



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

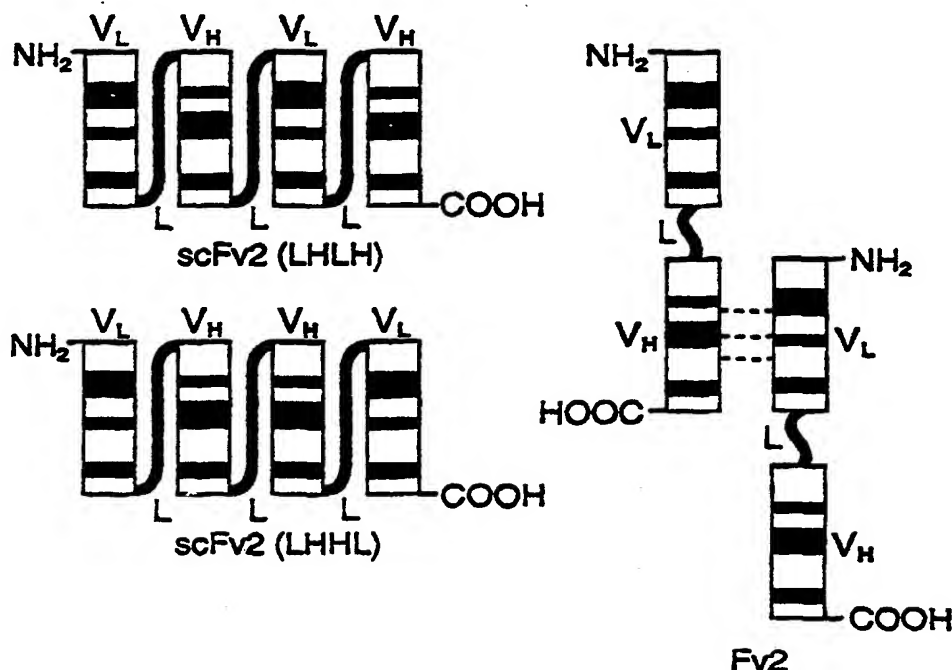
| | | |
|--|-----------|---|
| (51) International Patent Classification ⁵ : C12N 15/13, C07K 15/28, C12N 15/62, A61K 39/395 | A1 | (11) International Publication Number: WO 94/13806 |
| | | (43) International Publication Date: 23 June 1994 (23.06.94) |
| (21) International Application Number: PCT/US93/12039 | | (81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i> |
| (22) International Filing Date: 10 December 1993 (10.12.93) | | |
| (30) Priority Data: 07/990,263 11 December 1992 (11.12.92) US | | |
| (71) Applicant: THE DOW CHEMICAL COMPANY [US/US]; 2030 Dow Center, Abbott Road, Midland, MI 48640 (US). | | |
| (72) Inventors: MEZES, Peter, S.; 25 Sill Lane, Oldlyme, CT 06371 (US). GOURLIE, Brian, B.; 3713 Orchard Drive, Midland, MI 48640 (US). | | |
| (74) Agent: ULMER, Duane, C.; The Dow Chemical Company, Patent Department, P.O. Box 1967, Midland, MI 48641-1967 (US). | | |

(54) Title: MULTIVALENT SINGLE CHAIN ANTIBODIES

(57) Abstract

The present invention discloses multivalent single chain antibodies which have two or more biologically active antigen binding sites. The multivalent single chain antibodies are formed by using a peptide linker to covalently link two or more single chain antibodies, each single chain antibody having a variable light domain linked to a variable heavy chain domain by a peptide linker.

Schematic Representation Of Covalently & Non-Covalently Linked Single Chain Fv Multimers



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|----|--------------------------|----|---------------------------------------|----|--------------------------|
| AT | Austria | GB | United Kingdom | MR | Mauritania |
| AU | Australia | GE | Georgia | MW | Malawi |
| BB | Barbados | CN | Guinea | NE | Niger |
| BE | Belgium | GR | Greece | NL | Netherlands |
| BF | Burkina Faso | HU | Hungary | NO | Norway |
| BG | Bulgaria | IE | Ireland | NZ | New Zealand |
| BJ | Benin | IT | Italy | PL | Poland |
| BR | Brazil | JP | Japan | PT | Portugal |
| BY | Belarus | KE | Kenya | RO | Romania |
| CA | Canada | KG | Kyrgyzstan | RU | Russian Federation |
| CF | Central African Republic | KP | Democratic People's Republic of Korea | SD | Sudan |
| CG | Congo | KR | Republic of Korea | SE | Sweden |
| CH | Switzerland | KZ | Kazakhstan | SI | Slovenia |
| CI | Côte d'Ivoire | LI | Liechtenstein | SK | Slovakia |
| CM | Cameroon | LK | Sri Lanka | SN | Senegal |
| CN | China | LU | Luxembourg | TD | Chad |
| CS | Czechoslovakia | LV | Latvia | TG | Togo |
| CZ | Czech Republic | MC | Monaco | TJ | Tajikistan |
| DE | Germany | MD | Republic of Moldova | TT | Trinidad and Tobago |
| DK | Denmark | MG | Madagascar | UA | Ukraine |
| ES | Spain | ML | Mali | US | United States of America |
| FI | Finland | MN | Mongolia | UZ | Uzbekistan |
| FR | France | | | VN | Viet Nam |
| GA | Gabon | | | | |

MULTIVALENT SINGLE CHAIN ANTIBODIES

The present invention relates to single chain multivalent antibodies.

Antibodies are proteins belonging to a group of immunoglobulins elicited by the immune system in response to a specific antigen or substance which the body deems foreign. There are five classes of human antibodies, each class having the same basic structure. The basic structure of an antibody is a tetramer, or a multiple thereof, composed of two identical heterodimers each consisting of a light and a heavy chain. The light chain is composed of one variable (V) and one constant (C) domain, while a heavy chain is composed of one variable and three or more constant domains. The variable domains from both the light and heavy chain, designated V_L and V_H respectively, determine the specificity of an immunoglobulin, while the constant (C) domains carry out various effector functions.

Amino acid sequence data indicate that each variable domain comprises three complementarity determining regions (CDR) flanked by four relatively conserved framework regions (FR). The FR are thought to maintain the structural integrity of the variable region domain. The CDR have been assumed to be responsible for the binding specificity of individual antibodies and to account for the diversity of binding of antibodies.

As the basic structure of an antibody contains two heterodimers, antibodies are multivalent molecules. For example, the IgG classes have two identical antigen binding sites, while the pentameric IgM class has 10 identical binding sites.

Monoclonal antibodies having identical genetic parentage and binding specificity have been useful both as diagnostic and therapeutic agents. Monoclonal antibodies are routinely produced by hybridomas generated by fusion of mouse lymphoid cells with an appropriate mouse myeloma cell line according to established procedures. The administration of murine antibodies for in vivo therapy and diagnostics in humans is limited however, due to the human anti-mouse antibody response elicited by the human immune system.

Chimeric antibodies, in which the binding or variable regions of antibodies derived from one species are combined with the constant regions of antibodies derived from a different species, have been produced by recombinant DNA methodology. See, for example, Sahagen et al., *J. Immunol.*, 137:1066-1074 (1986); Sun et al., *Proc. Natl. Acad. Sci. USA*, 82:214-218 (1987); Nishimura et al., *Cancer Res.*, 47:999-1005 (1987); and Lie et al. *Proc Natl. Acad. Sci. USA*, 84:3439-3443 (1987) which disclose chimeric antibodies to tumor-associated antigens. Typically, the variable region of a murine antibody is joined with the constant region of a human antibody. It is expected that as such chimeric antibodies are largely human in composition, they will be substantially less immunogenic than murine antibodies.

Chimeric antibodies still carry the Fc regions which are not necessary for antigen binding, but constitute a major portion of the overall antibody structure which affects its pharmacokinetics. For the use of antibodies in immunotherapy or immunodiagnosics, is it

desirable to have antibody-like molecules which localize and bind to the target tissue rapidly and for the unbound material to quickly clear from the body. Generally, smaller antibody fragments have greater capillary permeability and are more rapidly cleared from the body than whole antibodies.

5 Since it is the variable regions of light and heavy chains that interact with an antigen, single chain antibody fragments (scFvs) have been created with one V_L and one V_H , containing all six CDR's, joined by a peptide linker (U.S. Patent 4,946,778) to create a V_L -L- V_H polypeptide, wherein the L stands for the peptide linker. A scFv wherein the V_L and V_H domains are orientated V_H -L- V_L is disclosed in U.S. Patent 5,132,405.

10 As the scFvs have one binding site as compared to the minimum of two for complete antibodies, the scFvs have reduced avidity as compared to the antibody containing two or more binding sites.

 It would therefore be advantageous to obtain constructions of scFvs having more than one binding site to enhance the avidity of the polypeptide, and retain or increase their
15 antigen recognition properties. In addition, it would be beneficial to obtain multivalent scFvs which are bispecific to allow for recognition of different epitopes on the target tissue, to allow for antibody-based recruitment of other immune effector functions, or allow antibody capture of a therapeutic or diagnostic moiety.

 It has been found that single chain antibody fragments, each having one V_H and
20 one V_L domain covalently linked by a first peptide linker, can be covalently linked by a second peptide linker to form a multivalent single chain antibody which maintains the binding affinity of a whole antibody. In one embodiment, the present invention is a multivalent single chain antibody having affinity for an antigen wherein the multivalent single chain antibody comprises two or more light chain variable domains and two or more heavy chain variable
25 domains; wherein, each variable domain is linked to at least one other variable domain.

 In another embodiment, the present invention is a multivalent single chain antibody which comprises two or more single chain antibody fragments, each fragment having affinity for an antigen wherein the fragments are covalently linked by a first peptide linker and each fragment comprising:

- 30 (a) a first polypeptide comprising a light chain variable domain;
 (b) a second polypeptide comprising a heavy chain variable domain; and
 (c) a second peptide linker linking the first and second polypeptides into a functional binding moiety.

 In another embodiment, the invention provides a DNA sequence which codes for
35 a multivalent single chain antibody, the multivalent single chain antibody comprising two or more single chain antibody fragments, each fragment having affinity for an antigen wherein the fragments are covalently linked by a first peptide linker and each fragment comprising:

- (a) a first polypeptide comprising a light chain variable domain;

- (b) a second polypeptide comprising a heavy chain variable domain; and
- (c) a second peptide linker linking the first and second polypeptides into a functional binding moiety.

The multivalent single chain antibodies allow for the construction of an antibody fragment which has the specificity and avidity of a whole antibody but are smaller in size allowing for more rapid capillary permeability. Multivalent single chain antibodies also allow for the construction of a multivalent single chain antibody wherein the binding sites can be two different antigenic determinants.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates covalently linked single chain antibodies having the configuration $V_L-L-V_H-L-V_L-L-V_H$ (LHLH) and $V_L-L-V_H-L-V_H-L-V_L$ (LHHL) and a noncovalently linked Fv single chain antibody (Fv2).

Figure 2 illustrates the nucleotide sequence of CC49 V_L .

Figure 3 illustrates the amino acid sequence of CC49 V_L .

Figure 4 illustrates the nucleotide sequence of CC49 V_H .

Figure 5 illustrates the amino acid sequence of CC49 V_H .

Figure 6 illustrates the nucleotide sequence and amino acid sequence of the CC49 single chain antibody LHLH in p49LHLH.

Figure 7 illustrates the nucleotide sequence and amino acid sequence of the CC49 single antibody LHHL in p49LHHL.

Figure 8 illustrates construction of plasmids pSL301 T and pSL301 HT.

Figure 9 illustrates construction of plasmid p49LHHL.

Figure 10 illustrates construction of plasmid p49LHLH.

Figure 11 illustrates the results of a competition assay using CC49 IgG, CC49 scFv2, and CC49 scFv using biotinylated CC49 IgG as competitor.

The entire teaching of all references cited herein are hereby incorporated by reference.

Nucleic acids, amino acids, peptides, protective groups, active groups and such, when abbreviated, are abbreviated according to the IUPAC IUB (Commission on Biological Nomenclature) or the practice in the fields concerned.

The term "single chain antibody fragment" (scFv) or "antibody fragment" as used herein means a polypeptide containing a V_L domain linked to a V_H domain by a peptide linker (L), represented by V_L-L-V_H . The order of the V_L and V_H domains can be reversed to obtain polypeptides represented as V_H-L-V_L . "Domain" is a segment of protein that assumes a discrete function, such as antigen binding or antigen recognition.

A "multivalent single chain antibody" means two or more single chain antibody fragments covalently linked by a peptide linker. The antibody fragments can be joined to form bivalent single chain antibodies having the order of the V_L and V_H domains as follows:

$V_L-L-V_H-L-V_L-L-V_H$; $V_L-L-V_H-L-V_H-L-V_L$; $V_H-L-V_L-L-V_H-L-V_L$; or $V_H-L-V_L-L-V_L-L-V_H$.

Single chain multivalent antibodies which are trivalent and greater have one or more antibody fragments joined to a bivalent single chain antibody by an additional interpeptide linker. In a preferred embodiment, the number of V_L and V_H domains is equivalent.

5 The present invention also provides for multivalent single chain antibodies which can be designated $V_H-L-V_H-L-V_L-L-V_L$ or $V_L-L-V_L-L-V_H-L-V_H$.

Covalently linked single chain antibodies having the configuration $V_L-L-V_H-L-V_L-L-V_H$ (LHLH) and $V_L-L-V_H-L-V_H-L-V_L$ (LHHL) are illustrated in Figure 1. A noncovalently linked Fv single chain antibody (Fv2) is also illustrated in Figure 1.

10 The single chain antibody fragments for use in the present invention can be derived from the light and/or heavy chain variable domains of any antibody. Preferably, the light and heavy chain variable domains are specific for the same antigen. The individual antibody fragments which are joined to form a multivalent single chain antibody may be directed against the same antigen or can be directed against different antigens.

15 To prepare a vector containing the DNA sequence for a single chain multivalent antibody, a source of the genes encoding for these regions is required. The appropriate DNA sequence can be obtained from published sources or can be obtained by standard procedures known in the art. For example, Kabat et al., *Sequences of Proteins of Immunological Interest 4th ed.* (1991), published by The U.S. Department of Health and Human Services, discloses
20 sequences of most of the antibody variable regions which have been described to date.

When the genetic sequence is unknown, it is generally possible to utilize cDNA sequences obtained from mRNA by reverse transcriptase mediated synthesis as a source of DNA to clone into a vector. For antibodies, the source of mRNA can be obtained from a wide range of hybridomas. See, for example, the catalogue ATCC Cell Lines and Hybridomas, American
25 Type Culture Collection, 20309 Parklawn Drive, Rockville Md., USA (1990). Hybridomas secreting monoclonal antibodies reactive with a wide variety of antigens are listed therein, are available from the collection, and usable in the present invention. These cell lines and others of similar nature can be utilized as a source of mRNA coding for the variable domains or to obtain antibody protein to determine amino acid sequence of the monoclonal antibody itself.

30 Variable regions of antibodies can also be derived by immunizing an appropriate vertebrate, normally a domestic animal, and most conveniently a mouse. The immunogen will be the antigen of interest, or where a hapten, an antigenic conjugate of the hapten to an antigen such as keyhole limpet hemocyanin (KLH). The immunization may be carried out conventionally with one or more repeated injections of the immunogen into the host mammal,
35 normally at two to three week intervals. Usually, three days after the last challenge, the spleen is removed and dissociated into single cells to be used for cell fusion to provide hybridomas from which mRNA can readily be obtained by standard procedures known in the art.

When an antibody of interest is obtained, and only its amino acid sequence is known, it is possible to reverse translate the sequence.

The V_L and V_H domains for use in the present invention are preferably obtained from one of a series of CC antibodies against tumor-associated glycoprotein 72 antigen (TAG-72) disclosed in published PCT Application WO 90/04410 on May 3, 1990, and published PCT Application WO 89/00692 on January 26, 1989. More preferred are the V_L and V_H domains from the monoclonal antibody designated CC49 in PCT Publications WO 90/04410 and WO 89/00692. The nucleotide sequence (SEQ ID NO: 1) which codes for the V_L of CC49 is substantially the same as that given in Figure 1. The amino acid sequence (SEQ ID NO: 2) of the V_L of CC49 is substantially the same as that given in Figure 2. The nucleotide sequence (SEQ ID NO: 3) which codes for the V_H of CC49 is substantially the same as that given in Figure 3. The amino acid sequence (SEQ ID NO: 4) for the V_H of CC49 is substantially the same as that given in Figure 4.

To form the antibody fragments and multivalent single chain antibodies of the present invention, it is necessary to have a suitable peptide linker. Suitable linkers for joining the V_H and V_L domains are those which allow the V_H and V_L domains to fold into a single polypeptide chain which will have a three dimensional structure very similar to the original structure of a whole antibody and thus maintain the binding specificity of the whole antibody from which antibody fragment is derived. Suitable linkers for linking the scFvs are those which allow the linking of two or more scFvs such that the V_H and V_L domains of each immunoglobulin fragment have a three dimensional structure such that each fragment maintains the binding specificity of the whole antibody from which the immunoglobulin fragment is derived. Linkers having the desired properties can be obtained by the method disclosed in U.S. Patent 4,946,778, the disclosure of which is hereby incorporated by reference. From the polypeptide sequences generated by the methods described in the 4,946,778, genetic sequences coding for the polypeptide can be obtained.

Preferably, the peptide linker joining the V_H and V_L domains to form a scFv and the peptide linker joining two or more scFvs to form a multivalent single chain antibody have substantially the same amino acid sequence.

It is also necessary that the linker peptides be attached to the antibody fragments such that the binding of the linker to the individual antibody fragments does not interfere with the binding capacity of the antigen recognition site.

A preferred linker is based on the helical linker designated 205C as disclosed in Pantoliano et al. *Biochem.*, 30, 10117-10125 (1991) but with the first and last amino acids changed because of the codon dictated by the Xho I site at one end and the Hind III site at the other. The amino acid sequence (SEQ ID NO: 5) of the preferred linker is as follows:

Leu-Ser-Ala-Asp-Asp-Ala-Lys-Lys-Asp-Ala-Ala-Lys-Lys-Asp-Asp-Ala-Lys-Lys-Asp-Asp-Ala-Lys-Lys-Asp-Leu.

The linker is generally 10 to 50 amino acid residues. Preferably, the linker is 10 to 30 amino acid residues. More preferably the linker is 12 to 30 amino acid residues. Most preferred is a linker of 15 to 25 amino acid residues.

Expression vehicles for production of the molecules of the invention include
5 plasmids or other vectors. In general, such vectors contain replicon and control sequences which are derived from species compatible with a host cell. The vector ordinarily carries a replicon site, as well as specific genes which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is readily transformed using pBR322 [Bolivar et al., *Gene*,
2, 95- (1977), or Sambrook et al., *Molecular Cloning*, Cold Spring Harbor Press, New York, 2nd
10 Ed. (1989)].

Plasmids suitable for eukaryotic cells may also be used. *S. cerevisiae*, or common baker's yeast, is the most commonly used among eukaryotic microorganisms, although a number of other strains, such as *Pichia pastoris*, are available. Cultures of cells derived from multicellular organisms such as SP2/0 or Chinese Hamster Ovary (CHO), which are available from
15 the ATCC, may also be used as hosts. Typical of vector plasmids suitable for mammalian cells are pSV2neo and pSV2gpt (ATCC); pSVL and pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnology, Inc.).

The use of prokaryotic and eukaryotic viral expression vectors to express the genes for polypeptides of the present invention is also contemplated.

20 It is preferred that the expression vectors and the inserts which code for the single chain multivalent antibodies have compatible restriction sites at the insertion junctions and that those restriction sites are unique to the areas of insertion. Both vector and insert are treated with restriction endonucleases and then ligated by any of a variety of methods such as those described in Sambrook et al., *supra*.

25 Preferred genetic constructions of vectors for production of single chain multivalent antibodies of the present invention are those which contain a constitutively active transcriptional promoter, a region encoding signal peptide which will direct synthesis/secretion of the nascent single chain polypeptide out of the cell. Preferably, the expression rate is commensurate with the transport, folding and assembly steps to avoid accumulation of the
30 polypeptide as insoluble material. In addition to the replicon and control sequences, additional elements may also be needed for optimal synthesis of single chain polypeptide. These elements may include splice signals, as well as transcription promoter, enhancers, and termination signals. Furthermore, additional genes and their products may be required to facilitate assembly and folding (chaperones).

35 Vectors which are commercially available can easily be altered to meet the above criteria for a vector. Such alterations are easily performed by those of ordinary skill in the art in light of the available literature and the teachings herein.

Additionally, it is preferred that the cloning vector contain a selectable marker, such as a drug resistance marker or other marker which causes expression of a selectable trait by the host cell. "Host cell" refers to cells which can be recombinantly transformed with vectors constructed using recombinant DNA techniques. A drug resistance or other selectable marker is intended in part to facilitate in the selection of transformants. Additionally, the presence of a selectable marker, such as a drug resistance marker, may be of use in keeping contaminating microorganisms from multiplying in the culture medium. In this embodiment, such a pure culture of the transformed host cell would be obtained by culturing the cells under conditions which require the induced phenotype for survival.

Recovery and purification of the present invention can be accomplished using standard techniques known in the art. For example, if they are secreted into the culture medium, the single chain multivalent antibodies can be concentrated by ultrafiltration. When the polypeptides are transported to the periplasmic space of a host cell, purification can be accomplished by osmotically shocking the cells, and proceeding with ultrafiltration, antigen affinity column chromatography or column chromatography using ion exchange chromatography and gel filtration. Polypeptides which are insoluble and present as refractile bodies, also called inclusion bodies, can be purified by lysis of the cells, repeated centrifugation and washing to isolate the inclusion bodies, solubilization, such as with guanidine-HCl, and refolding followed by purification of the biologically active molecules.

The activity of single chain multivalent antibodies can be measured by standard assays known in the art, for example competition assays, enzyme-linked immunosorbant assay (ELISA), and radioimmunoassay (RIA).

The multivalent single chain antibodies of the present invention provide unique benefits for use in diagnostics and therapeutics. The use of multivalent single chain antibodies afford a number of advantages over the use of larger fragments or entire antibody molecules. They reach their target tissue more rapidly, and are cleared more quickly from the body.

For diagnostic and/or therapeutic uses, the multivalent single chain antibodies can be constructed such that one or more antibody fragments are directed against a target tissue and one or more antibody fragments are directed against a diagnostic or therapeutic agent.

The invention also concerns pharmaceutical compositions which are particularly advantageous for use in the diagnosis and/or therapy of diseases, such as cancer, where target antigens are often expressed on the surface of cells. For diagnostic and/or therapeutic uses, the multivalent single chain antibodies can be conjugated with an appropriate imaging or therapeutic agent by methods known in the art. The pharmaceutical compositions of the invention are prepared by methods known in the art, e.g., by conventional mixing, dissolving or lyophilizing processes.

The invention will be further clarified by a consideration of the following examples, which are intended to be purely exemplary of the present invention.

5

10

15

20

25

30

35

ABBREVIATIONS

| | | |
|----|------------------|---|
| | BCIP | 5-bromo-4-chloro-3-indoyl phosphate |
| | bp | base pair |
| 5 | Bis-Tris propane | (1,3-bis[tris(hydroxymethyl)-methylamino]-propane) |
| | BSA | bovine serum albumin |
| | CDR | Complementarity determining region |
| | ELISA | enzyme linked immunosorbent assay |
| 10 | Fv2 | non-covalent single chain Fv dimer |
| | IEF | isoelectric focusing |
| | Kbp | kilo base pair |
| | LB | Luria-Bertani medium |
| | Mab | monoclonal antibody |
| 15 | MES | 2-(N-Morpholino)ethane sulfonic acid |
| | MW | molecular weight |
| | NBT | nitro blue tetrazolium chloride |
| | Oligo | Oligonucleotides |
| | PAG | polyacrylamide gel |
| 20 | PAGE | polyacrylamide gel electrophoresis |
| | PBS | phosphate buffered saline |
| | PCR | polymerase chain reaction |
| | pSCFV | plasmid containing DNA sequence coding for SCFV |
| | RIGS | radioimmunoguided surgery |
| 25 | RIT | radioimmunotherapy |
| | scFv | single chain Fv immunoglobulin fragment monomer |
| | scFv2 | single chain Fv immunoglobulin fragment dimer covalently linked |
| | SDS | sodium dodecyl sulfate |
| | TBS | Tris-buffered saline |
| 30 | Tris | (Tris[hydroxymethyl]aminomethane) |
| | TTBS | Tween-20 wash solution |
| | V _H | immunoglobulin heavy chain variable domain |
| | V _L | immunoglobulin light chain variable domain |

35

Antibodies

CC49: A murine monoclonal antibody specific to the human tumor-associated glycoprotein 72 (TAG-72) deposited as ATCC No. HB9459.

5 CC49 FAB: An antigen binding portion of CC49 consisting of an intact light chain linked to the N-terminal portion of the heavy chain.

CC49 scFv: Single chain antibody fragment consisting of two variable domains of CC49 antibody joined by a peptide linker.

10 CC49 Fv2: Two CC49 scFv non-covalently linked to form a dimer. The number after Fv refers to the number of monomer subunits of a given molecule, e.g., CC49 Fv6 refers to the hexamer multimers.

CC49 scFv2: Covalently-linked single chain antibody fragment consisting of two CC49 V_L domains and two V_H domains joined by three linkers. Six possible combinations for the order of linking the V_L(L) and the V_H(H) domains together are: LHLH, LHHL, LLHH, HLLH, HLHL, and HHLL.

15 Plasmids

pSCFV UHM: Plasmid containing coding sequence for scFv consisting of a CC49 variable light chain and a CC49 variable heavy chain joined by a 25 amino acid linker.

p49LHLH or p49LHHL: Plasmids containing the coding sequence for producing CC49 scFv2 LHLH or LHHL products, respectively.

20 EXAMPLES

General Experimental

Procedures for molecular cloning are as those described in Sambrook et al., *Molecular Cloning*, Cold Spring Harbor Press, New York, 2nd Ed. (1989) and Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, New York (1992), the disclosures
25 of which are hereby incorporated by reference.

All water used throughout was deionized distilled water.

Oligonucleotide Synthesis and Purification

All oligonucleotides (oligos) were synthesized on either a Model 380A or a Model 391 DNA Synthesizer from Applied Biosystems (Foster City, CA) using standard
30 β -cyanoethyl phosphoramidites and synthesis columns. Protecting groups on the product were removed by heating in concentrated ammonium hydroxide at 55°C for 6 to 15 hours. The ammonium hydroxide was removed through evaporation and the crude mixtures were resuspended in 30 to 40 μ L of sterile water. After electrophoresis on polyacrylamide-urea gels, the oligos were visualized using short wavelength ultraviolet (UV) light. DNA bands were
35 excised from the gel and eluted into 1 mL of 100 mM Tris-HCl, pH 7.4, 500 mM NaCl, 5 mM EDTA over 2 hours at 65°C. Final purification was achieved by applying the DNA to Sep-Pac™ C-18 columns (Millipore, Bedford, MA) and eluting the bound oligos with 60 percent methanol. The

solution volume was reduced to approximately 50 μ L and the DNA concentration was determined by measuring the optical density at 260 nm (OD_{260}).

Restriction Enzyme Digests

All restriction enzyme digests were performed using Bethesda Research Laboratories (Gaithersburg, MD), New England Biolabs, Inc. (Beverly, MA) or Boehringer Mannheim (BM, Indianapolis, IN) enzymes and buffers following the manufacturer's recommended procedures. Digested products were separated by polyacrylamide gel electrophoresis (PAGE). The gels were stained with ethidium bromide, the DNA bands were visualized using long wavelength UV light and the DNA bands were then excised. The gel slices were placed in dialysis tubing (Union Carbide Corp., Chicago) containing 5 mM Tris, 2.5 mM acetic acid, 1 mM EDTA, pH 8.0 and eluted using a Max Submarine electrophoresis apparatus (Hoefer Scientific Instruments, CA). Sample volumes were reduced on a Speed Vac Concentrator (Savant Instruments, Inc., NY). The DNA was ethanol precipitated and redissolved in sterile water.

Enzyme Linked Immunosorbent Assay (ELISA)

TAG-72 antigen, prepared substantially as described by Johnson et al, *Can. Res.*, 46, 850-857 (1986), was adsorbed onto the wells of a polyvinyl chloride 96 well microtiter plate (Dynatech Laboratories, Inc., Chantilly, VA) by drying overnight. The plate was blocked with 1 percent BSA in PBS for 1 hour at 31°C and then washed 3 times with 200 μ L of PBS, 0.05 percent Tween-20. 25 μ L of test antibodies and 25 μ L of biotinylated CC49 (1/20,000 dilution of a 1 mg/mL solution) were added to the wells and the plate incubated for 30 minutes at 31°C. The relative amounts of TAG-72 bound to the plate, biotinylated CC49, streptavidin-alkaline phosphatase, and color development times were determined empirically in order not to have excess of either antigen or biotinylated CC49, yet have enough signal to detect competition by scFv. Positive controls were CC49 at 5 μ g/mL and CC49 Fab at 10 μ g/mL. Negative controls were 1 percent BSA in PBS and/or concentrated LB. Unbound proteins were washed away. 50 μ L of a 1:1000 dilution of streptavidin conjugated with alkaline phosphatase (Southern Biotechnology Associates, Inc., Birmingham, AL) were added and the plate was incubated for 30 minutes at 31°C. The plate was washed 3 more times. 50 μ L of a para-nitrophenyl-phosphate solution (Kirkegaard & Perry Laboratories, Inc., Gaithersburg, MD) were added and the color reaction was allowed to develop for a minimum of 20 minutes. The relative amount of scFv2 binding was measured by optical density scanning at 404-450 nm using a microplate reader (Molecular Devices Corporation, Manlo Park, CA). Binding of the scFv2 species resulted in decreased binding of the biotinylated CC49 with a concomitant decrease in color development.

SDS-PAGE and Western Blotting

Samples for SDS-PAGE analysis (20 μ L) were prepared by boiling in a non-reducing sample preparation buffer-Sepasol I (Integrated Separation Systems (ISS), Natick, MA) for

5 minutes and loaded on 10-20 percent gradient polyacrylamide Daiichi Minigels as per the manufacturer's directions (ISS).

Electrophoresis was conducted using a Mini 2-gel apparatus (ISS) at 55 mA per gel at constant current for approximately 75 minutes. Gels were stained in Coomassie Brilliant Blue R-250 (Bio-Rad, Richmond, CA) for at least 1 hour and destained. Molecular weight standards were prestained (Mid Range Kit, Diversified Biotech, Newton Center, MA) and included the following proteins: Phosphorylase b, glutamate dehydrogenase, ovalbumin, lactate dehydrogenase, carbonic anhydrase, B-lactoglobulin and cytochrome C. The corresponding MWs are: 95,500, 55,000, 43,000, 36,000, 29,000, 18,400, and 12,400, respectively.

When Western analyses were conducted, a duplicate gel was also run. After electrophoresis, one of the gels was equilibrated for 15-20 minutes in anode buffer #1 (0.3 M Tris-HCl pH 10.4). An Immobilon-P PVDF (polyvinylidene dichloride) membrane (Millipore, Bedford, MA) was treated with methanol for 2 seconds, and immersed in water for 2 minutes. The membrane was then equilibrated in anode buffer #1 for 3 minutes. A Milliblot-SDE apparatus (Millipore) was utilized to transfer proteins in the gel to the membrane. A drop of anode buffer #1 was placed in the middle of the anode electrode surface. A sheet of Whatman 3MM filter paper was soaked in anode buffer #1 and smoothly placed on the electrode surface. Another filter paper soaked in anode buffer #2 (25 mM tris pH 10.4) was placed on top of the first one. A sandwich was made by next adding the wetted PVDF membrane, placing the equilibrated gel on top of this and finally adding a sheet of filter paper soaked in cathode buffer (25mM Tris-HCl, pH 9.4 in 40 mM glycine). Transfer was accomplished in 30 minutes using 250 mA constant current (initial voltage ranged from 8-20 volts).

After blotting, the membrane was rinsed briefly in water and placed in a dish with 20 mL blocking solution (1 percent bovine serum albumin (BSA) (Sigma, St. Louis, MO) in Tris-buffered saline (TBS)). TBS was purchased from Pierce Chemical (Rockford, IL) as a preweighed powder such that when 500 mL water is added, the mixture gives a 25 mM Tris, 0.15 M sodium chloride solution at pH 7.6. The membranes were blocked for a minimum of 1 hour at ambient temperature and then washed 3 times for 5 minutes each using 20 mL 0.5 percent Tween-20 wash solution (TTBS). To prepare the TTBS, 0.5mL of Tween 20 (Sigma) was mixed per liter of TBS. The probe antibody used was 20 mL biotinylated FAID14 solution (10 µg per 20 mL antibody buffer). Antibody buffer was made by adding 1 g BSA per 100 mL of TTBS. After probing for 30-60 minutes at ambient temperature, the membrane was washed 3 times with TTBS, as above.

Next, the membrane was incubated for 30-60 minutes at ambient temperature with 20 mL of a 1:500 dilution in antibody buffer of streptavidin conjugated with alkaline phosphatase (Southern Biotechnology Associates, Birmingham, AL). The wash step was again repeated after this, as above. Prior to the color reaction, membranes were washed for 2 minutes in an alkaline carbonate buffer (20 mL). This buffer is 0.1 M sodium bicarbonate,

1 mM $\text{MgCl}_2 \cdot \text{H}_2\text{O}$, pH 9.8. To make up the substrate for alkaline phosphatase, nitroblue tetrazolium (NBT) chloride (50 mg, Sigma) was dissolved in 70 percent dimethylformamide. 5-Bromo-4-chloro-3-indoyl phosphate (BCIP) (25 mg, Sigma) was separately dissolved in 100 percent dimethylformamide. 5-Bromo-4-chloro-3-indoyl phosphate (BCIP) 25 mg, Sigma) was separately dissolved in 100 percent dimethylformamide. These solutions are also commercially available as a Western developing agent sold by Promega. For color development, 120 μL of each were added to the alkaline solution above and allowed to react for 15 minutes before they were washed from the developed membranes with water.

Biotinylated FAID14

FAID14 is a murine anti-idiotypic antibody (IgG2a, K isotype) deposited as ATCC No. CRL 10256 directed against CC49. FAID14 was purified using a Nygene Protein A affinity column (Yonkers, NY). The manufacturer's protocol was followed, except that 0.1 M sodium citrate, pH 3.0 was used as the elution buffer. Fractions were neutralized to pH ~ 7 using 1.0 M Tris-HCl pH 9.0. The biotinylation reaction was set up as follows. FAID14 (1 mg, 100 μL in water) was mixed with 100 μL of 0.1 M Na_2CO_3 , pH 9.6. Biotinyl- ϵ -amino-caproic acid N-hydroxy succinimide ester (Biotin-X-NHS) (Calbiochem, LaJolla, CA) (2.5 mg) was dissolved in 0.5 mL dimethylsulfoxide. Biotin-X-NHS solution (20 μL) was added to the FAID14 solution and allowed to react at 22°C for 4 hours. Excess biotin and impurities were removed by gel filtration, using a Pharmacia Superose 12 HR10/30 column (Piscataway, NJ). At a flow rate of 0.8 mL/min, the biotinylated FAID14 emerged with a peak at 16.8 min. The fractions making up this peak were pooled and stored at 4°C and used to detect the CC49 idotype as determined by the CC49 V_L and V_H CDRs.

Isoelectric Focusing (IEF)

Isoelectric points (pI's) were predicted using a computer program called PROTEIN-TITRATE, available through DNASTAR (Madison, WI). Based on amino acid composition with an input sequence, a MW value is given, in addition to the pI. Since Cys residues contribute to the charge, the count was adjusted to 0 for Cys, since they are all involved in disulfide bonds.

Experimentally, pI's were determined using Isogel agarose IEF plates, pH range 3-10 (FMC Bioproducts, Rockland, ME). A Biorad Bio-phoresis horizontal electrophoresis cell was used to run the IEF, following the directions of both manufacturers. The electrophoresis conditions were: 500 volts (limiting), at 20 mA current and 10 W of constant power. Focusing was complete in 90 min. IEF standards were purchased from Biorad; the kit included phycocyanin, β -lactoglobulin B, bovine carbonic anhydrase, human carbonic anhydrase, equine myoglobin, human hemoglobins A and C, 3 lentil lectins and cytochrome C, with pI values of 4.65, 5.10, 6.00, 6.50, 7.00, 7.10 and 7.50, 7.80, 8.00, and 8.20 and 9.60, respectively. Gels were stained and destained according to the directions provided by FMC.

Quantitation of CC49 Antibody Species

All purified CC49 antibodies including the IgG, scFv2 species and the monomeric scFv were quantitated by measuring the absorbance of protein dilutions at 280 nm using matching 1.0 cm pathlength quartz cuvettes (Hellma) and a Perkin-Elmer UV/VIS Spectrophotometer, Model 552A. Molar absorptivities (E_m) were determined for each antibody by using the following formula:

$$E_m = (\text{number Trp}) \times 5,500 + (\text{number Tyr}) \times 1,340 + (\text{number (Cys)}_2) \times 150 + (\text{number Phe}) \times 10$$

The values are based on information given by D. B. Wetlaufer, *Advances in Protein Chemistry*, 17, 375-378).

High Performance Liquid Chromatography

All high performance liquid chromatography (HPLC) was performed for CC49 scFv2 purification using an LKB HPLC system with titanium or teflon tubing throughout. The system consists of the Model 2150 HPLC pump, model 2152 controller, UV CORD SII model 2238 detection system set at an absorbance of 276 nm and the model 2211 SuperRac fraction collector.

PCR Generation of Subunits

All polymerase chain reactions (PCR) were performed with a reaction mixture consisting of: 150 picograms (pg) plasmid target (pSCFVUHM); 100 pmoles primers; 1 μ L Perkin-Elmer-Cetus (PEC, Norwalk, CT) Ampli-Taq polymerase; 16 μ L of 10 mM dNTPs and 10 μ L of 10X buffer both supplied in the PEC kit; and sufficient water to bring the volume to total volume to 100 μ L. The PCR reactions were carried out essentially as described by the manufacturer. Reactions were done in a PEC 9600 thermocycler with 30 cycles of: denaturation of the DNA at 94°C for 20 to 45 sec, annealing from between 52 to 60°C for 0.5 to 1.5 min., and elongation at 72°C for 0.5 to 2.0 min. Oligonucleotide primers were synthesized on an Applied Biosystems (Foster City, CA) 380A or 391 DNA synthesizer and purified as above.

Ligations

Ligation reactions using 100 ng of vector DNA and a corresponding 1:1 stoichiometric equivalent of insert DNA were performed using a Stratagene (La Jolla, CA) T4 DNA ligase kit following the manufacturer's directions. Ligation reactions (20 μ L total volume) were initially incubated at 18°C and allowed to cool gradually overnight to 4°C.

Transformations

Transformations were performed utilizing 100 μ L of Stratagene E. coli AG1 competent cells (Stratagene, La Jolla, CA) according to the directions provided by the manufacturer. DNA from the ligation reactions (1-5 μ L) were used. After the transformation step, cells were allowed to recover for 1 hr in Luria broth (LB) at 37°C with continuous mixing and subsequently plated onto either 20 μ g/mL chloramphenicol containing (CAM 20) Luria agar for pSCFVUHM, p49LHLH or p49LHHL or 100 μ g/mL ampicillin (AMP 100) Luria agar plates

(LB-AMP 100) for clones containing the plasmid pSL301 or subsequent constructions derived from pSL301.

Screening of E. coli Clones

5 Bacterial plasmids were isolated from LB broth culture containing the appropriate drug to maintain selection pressure using Promega (Madison, WI) Magic mini-prep plasmid preparation kits. The kit was used per the manufacturer's specifications.

Plasmid Constructions

Two plasmids, designated p49LHLH and p49LHHL, were constructed to produce multivalent single chain antibodies. The host cell containing p49LHLH produced a polypeptide
10 which can be designated by V_L -L- V_H -L- V_L -L- V_H where V_L and V_H are the light and heavy chain variable regions of CC49 antibody and linker (L) is a 25 amino acid linker having the sequence (SEQ ID NO: 5).

Leu-Ser-Ala-Asp-Asp-Ala-Lys-Lys-Asp-Ala-Ala-Lys-Lys-Asp-Asp-Ala-Lys-Lys-Asp-
-Asp-Ala-Lys-Lys-Asp-Leu.

15 The host cell containing p49LHHL produced a polypeptide which can be designated by V_L -L- V_H -L- V_H -L- V_L where V_L and V_H are the light and heavy chain variable domains of the CC49 antibody and L is a peptide linker having the amino acid sequence indicated above.

The nucleotide sequence (SEQ ID NO: 6) and amino acid sequence (SEQ ID NO: 7) of the CC49 V_L -L- V_H -L- V_L -L- V_H (p49LHLH) are given in Figure 6. The nucleotide sequence (SEQ
20 ID NO: 8) and amino acid sequence (SEQ ID NO: 9) of the CC49 V_L -L- V_H -L- V_H -L- V_L (p49LHHL) are given in Figure 7.

Construction of pSL301 HT

The construction of pSL301 HT is illustrated in Figure 8. The *Bacillus licheniformis*
25 penicillinase P (penP) terminator sequence was removed from the plasmid designated pSCFV UHM by a 45 minute digest with Nhe I and BamH I, excised from a 4.5 percent polyacrylamide gel after electrophoresis, electroeluted, ethanol precipitated and ligated into the same sites in the similarly prepared vector: pSL301 (Invitrogen, San Diego, CA). A procedure for preparing pSCFV UHM is given in U.S. patent application Ser. No. 07/935,695 filed
30 August 21, 1992, the disclosure of which is hereby incorporated by reference. In general, pSCFV UHM contains a nucleotide sequence for a penP promoter; a unique Nco I restriction site; CC49 V_L region; Hind III restriction site; a 25 amino acid linker; a unique Xho I restriction site; CC49 V_H region; Nhe I restriction site; penP terminator; and BamH I restriction site (see, Figure 8). The penP promoter and terminator are described in Mezes, et al. (1983), *J. Biol.*
35 *Chem.*, 258, 11211-11218 (1983).

An aliquot of the ligation reaction (3 μ L) was used to transform competent *E. coli* AG1 cells which were plated on LB-AMP100 agar plates and grown overnight. Potential clones containing the penP terminator insert were screened using a Pharmacia (Gaithersburg, MD) T7

Quickprime ³²P DNA labeling kit in conjunction with the microwave colony lysis procedure outlined in Buluwela et al., *Nucleic Acid Research*, 17, 452 (1989). The probe, which was the penP-Nhe I-BamH I terminator fragment itself was prepared and used according to the directions supplied with the Quickprime kit. A clone which was probe positive and which contained the 207 base pair inserts from a BamH I and Nhe I digest (base pairs (bp) 1958 to 2165, Figure 6) was designated pSL301 T and chosen to construct pSL301 HT which would contain the nucleotide sequence for CC49 V_H. The reason the Nhe I-BamH I penP terminator was placed into pSL301 was to eliminate the Eco47 III restriction endonuclease site present in the polylinker region between its Nhe I and BamH I sites. This was designed to accommodate the subsequent build-up of the V_L and V_H domains where the Eco47 III site needed to be unique for the placement of each successive V domain into the construction. As each V domain was added at the Eco47 III-Nhe I sites, the Eco47 III was destroyed in each case to make the next Eco47 III site coming in on the unique insert.

The V_H sequence was made by PCR with oligos 5' SCP1 and 3' oligo SCP5 using pSCFV UHM as the target for PCR amplification. The DNA sequence for SCP1 (SEQ ID NO: 10) and SCP5 (SEQ ID NO: 11) are as follows:

SCP1: 5'-TAAA CTC GAG GTT CAG TTG CAG CAG -3'

SCP5: 5'-TAAA GCT AGC ACCA AGC GCT TAG TGA GGA GAC GGT GAC TGA GGT-3'

The underlined portion indicates the endonuclease restriction sites.

The amplified V_H DNA was purified from a 4 percent PAG, electroeluted ethanol precipitated and dissolved in 20 µL water. The V_H sequence was digested with Xho I and Nhe I restriction enzymes and used as the insert with the pSL301 T vector which had been digested with the same restriction enzymes and subsequently purified. A standard ligation reaction was done and an aliquot (4 µL) used to transform competent E. coli AG1 cells. The transformed cells were plated onto LB AMP100 agar plates. Candidate clones were picked from a Nhe I and Xho I digest screen that revealed that the CC49V_H insert had been obtained.

DNA sequencing was performed to verify the sequence of the CC49V_H with United States Biochemical (USB) (Cleveland, Ohio) Sequence kit and sequencing primers pSL301SEQB (a 21 bp sequencing primer which annealed in the pSL301 vector 57 bp upstream from the Xho I site) and CC49VHP, revealed clones with the correct CC49V_H sequence in pSL301HT. This plasmid was used as the starting point in the construction of both pSL301-HHLT and pSL301-HLHT. The sequencing oligos used are shown here.

The nucleotide sequence of pSL301SEQ B (SEQ ID NO: 12) and CC49V_H (SEQ ID No: 13) are as follows:

pSL301SEQB: 5'-TCG TCC GAT TAG GCA AGC TTA-3'

CC49VHP: 5'-GAT GAT TTT AAA TAC AAT GAG-3'

Example 1 p49LHHL Construction

Using pSL301 HT (5 µg) as the starting material, it was digested with Eco47 III and Nhe I and the larger vector fragment was purified. A CC49V_H insert fragment was generated by PCR using SCP6C as the 5' oligo and SCP5 as the 3' oligo. The nucleotide sequence (SEQ ID NO: 14) of SCP6B is as follows:

SCP6B: 5'-TAAA TGC GCA GAT GAC GCA AAG AAA GAC GCA GCT AAA AAA GAC GAT
GCC AAA AAG GAT GAC GCC AAG AAA GAT CTT GAG GTT CAG TTG CAG CAG
TCT-G'

The oligo SCP6B also contains part of the coding region for the linker (bp 8-76 of SEQ ID NO: 14). The portion of the oligo designed to anneal with the CC49VH target in pSCFV UHM is from bp77-90 in SEQ ID NO: 14.

The underlined sequence corresponds to the Fsp I site. The resulting PCR insert was purified, digested with Fsp I and Nhe I and used in a ligation reaction with the pSL301 HT Eco47 III-Nhe I vector (Figure 7). Competent E. coli AG1 cells were used for the transformation of this ligation reaction (3 µL) and were plated on LB-AMP100 agar plates. Two clones having the correct size Xho I-Nhe I insert representative of the pSL301 HHT product were sequenced with the oligo SQP1 and a single clone with the correct sequence (nucleotides 1124-1543 of Figure 7) was chosen for further construction. The nucleotide sequence of SQP1 (SEQ ID NO: 16) is as follows:

SQP1: 5'-TG ACT TTA TGT AAG ATG ATG T-3'

The final linker-V_L subunit (bp 1544-1963, Figure 7) was generated using the 5' oligo, SCP7b and the 3' oligo, SCP8a, using pSCFV UHM as the target for the PCR. The nucleotide sequence of SCP7b (SEQ ID NO: 17) is as follows:

SCP7b: 5'-TAAA TGC GCA GAT GAC GCA AAG AAA GAC GCA GCT AAA AAA GAC GAT
GCC AAA AAG GAT GAC GCC AAG AAA GAT CTT GAC ATT GTG ATG TCA CAG TCT
CC

The underlined nucleotides correspond to an Fsp I site. The nucleotide sequence of SCP8a (SEQ ID NO: 18) is as follows:

SCP8a: 5'-TAAA GCT AGC TTT TTA CTT AAG CAC CAG CTT GGT CCC-3'

The first set of underlined nucleotides correspond to an Nhe I site, while the other corresponds to an Afl II site. Nucleotides 8-76 of SCP70 code for the linker (nucleotides 1544-1612 of Figure 7) while nucleotides 77-99 which anneal to the V_L correspond to 1613-1635 of Figure 7. The primer SCP8a has a short tail at its 5' end, a Nhe I restriction site, a stop codon, an Afl II restriction site and the last 21 bases of the V_L. After Fsp I and Nhe I digestion, this resulting 420 bp insert was purified and ligated into the Nhe I and Eco47 III sites of the purified pSL301HHT vector, candidate clones were screened with Nhe I and Xho I, the correct size insert verified and sequenced with 49LFR2(-) and SQP1 to confirm the newly inserted sequence in pSL301HHLT. The nucleotide sequence (SEQ ID NO: 19) is as follows:

49LFR2(-): 5'-CTG CTG GTA CCA GGC CAA G-3'

The plasmid pSL301HHLT was digested with Xho I and Nhe I, purified, and the resulting 1179 bp V_H-linker-V_H-linker-V_L segment ligated into pSCFV UHM, which had been cut with the same restriction enzymes and the larger vector fragment purified, to form p49LHHL.

- 5 The ligation reaction (4 μ L aliquot) was used to transform competent *E. coli* AG1 cells (Stratagene) and plated onto LBCAM20 agar plates. A single clone which had a plasmid with the correct restriction enzyme map was selected to contain p49LHHL. The p49LHHL contains a penP promoter and a nucleotide sequence for the CC49 multivalent single chain antibody scFv2:

- 10 V_L-L-V_H-L-V_H-L-V_L or CC49 scFv2 (LHHL).

Example 2: p49LHLH Construction

The construction of p49LHLH is schematically represented in Figure 11. A linker-V_L subunit was generated with the 5' oligo SCP7b and the 3' oligo SCP9.

SCP9: 5'-TAA AGC TAG CAC CAA GCG CTT AGT TTC AGC ACC AGC TTG GTC CCA G-3'

- 15 The SCP7b oligo (nucleotides 8-76) codes for the linker in Figure 6 (corresponding to nucleotides 1124-1192) and annealed to the pSCFV UHM target for the PCR (nucleotides 77-99) corresponding to nucleotides 1193-1215 of the V_L in Figure 6.

- SCP9 has a Nhe I site (first underlined nucleotides) and an Eco47 III site (second underlined nucleotides) which are restriction sites needed for making the pSL301HHLT ready to accept the next V domain. Nucleotides 18-23 of SCP9 correspond to nucleotides 1532-1537 of Figure 6 (coding for the first 2 amino acids of the linker), while nucleotides 24-46 correspond to nucleotides 1508-1531 of Figure 6 which was also the annealing region for SCP9 in the PCR. The plasmid pSL301 HT was digested with Eco47 III and Nhe I and the larger vector fragment was purified for ligation with the linker-CC49V_L DNA insert fragment from the PCR which had been treated with Fsp I and Nhe I and purified. The ligation mixture (3 μ L) was used to transform *E. coli* AG1 competent cells and one colony having the correct Xho I-Nhe I size fragment was sequenced using the oligo PENPTSEQ2. The nucleotide sequence (SEQ. ID NO. 21) is as follows:

5'-TTG ATC ACC AAG TGA CTT TAT G-3'

- 30 The sequencing results indicated that there had been a PCR error and deletion in the resulting pSL301HT clone. A five base deletion, corresponding to nucleotides 1533-1537 as seen in Figure 6 had been obtained and nucleotide 1531 which should have been a T was actually a G, as determined from the DNA sequence data. The resulting sequence was

5'...G AAGC GCT T...etc.

- where the underlined sequence fortuitously formed an Eco47 III site. The AGCGCT sequence in Figure 6, would correspond to nucleotides 1530, 1531, 1532, 1538, 1539 and 1540. This error was corrected in the next step, generating pSL301 HLHT, by incorporating the 5 base deletion at the end of oligo SCP6C.

SCP6C: 5'-TAAGCGCTGATGATGCTAAGAAGGACGCCGCAAAAAA
 GGACGACGCAAAAAAAGATGATGCAAAAAAAGGATCTGG
 AGGTTCAAGTTGCAGCAGTCTGAC-3'

The underlined sequence in SCP6c corresponds to an Eco47 III site. SCP6C was
 5 used as the 5' oligo, with SCP10 as the 3' oligo in a PCR to generate a linker CC49 V_L segment.
 The nucleotide sequence (SEQ ID NO: 23) is as follows:

SCP10: 5'TTG TGC TAG CTT TTT ATG AGG AGA CGG TGA CTG AGG TT-3'

The underlined sequence in SCP10 corresponds to the Nhe I site found at
 nucleotides 1958-1963 in Figure 6. The PCR insert was digested this time only with Nhe I and
 10 purified. The vector (pSL301 HLT) was digested at the Eco47 III site (that had been formed) and
 Nhe I and purified. The insert and vector were ligated and an aliquot (3 µL) used to transform
 competent E. coli AG1 cells. This was plated on LB-AMP100 plates and candidate clones
 screened with Xho I and Nhe I. Three clones having the correct size DNA were obtained. Two
 of these clones were sequenced using the oligo 49VLCDR3(+) and SQP1. The nucleotide
 15 sequence (DWQ ID NO: 24 of 49VLCDR3(+)) is as follows:

49VLCDR3(+):

5'-CAG CAG TAT TAT AGC TAT-3'

One clone, with the correct sequence was obtained and the sequence from
 nucleotides 1533 to 1963 in Figure 6 were verified, giving a correct pSL301 HLHL clone.

To generate the final plasmid, p49LHLH for expression in E. coli, pSL301 HLHT
 20 (5 µg) was digested with Nhe I and Xho I, and the smaller insert fragment containing the
 V_H-L-V_L-L-V_H sequence purified. It was ligated with the larger purified vector fragment from a
 digest of pSCFV UHM (5 µg) with Xho I and Nhe I. An aliquot of the ligation mix (4 µL) was used
 to transform competent E. coli AG1 cells. The transformation mix was plated on LB-CAM20
 25 plates, and a representative clone for p49 LHLH was selected on the basis of a correct restriction
 enzyme map (see Figure 10) and biological activity toward TAG-72.

Example 3: Purification of CC49 scFv2 LHLH and LHLH Covalently Linked Dimers

For the purification of the CC49 covalently linked single chain dimers, (scFv2),
 E. coli periplasmic fractions were prepared from 1.0 L overnight cultures of both p49LHLH and
 30 p49LHHL. Briefly, the culture was divided into 4 X 250 mL portions and centrifuged at
 5,000 rpm for 10 minutes in a Sorvall GS-3 rotor. The pelleted cells were washed and
 resuspended in 100 mL each of 10 mM Tris-HCl pH 7.3 containing 30 mM NaCl. The cells were
 again pelleted and washed with a total of 100 mL 30 mM Tris-HCl pH 7.3 and pooled into one
 tube. To this, 100 mL of 30 mM Tris-HCl pH 7.3 containing 40 percent w/v sucrose and 2.0 mL of
 35 10 mM EDTA pH 7.5 was added. The mixture was kept at room temperature, with occasional
 shaking, for 10 minutes. The hypertonic cells were then pelleted as before. In the next step, the
 shock, the pellet was quickly suspended in 20 mL ice cold 0.5 mM MgCl₂ and kept on ice for 10
 minutes, with occasional shaking. The cells were pelleted as before and the supernatant

containing the *E. coli* periplasmic fraction was clarified further by filtration through a 0.2 μ m Nalge (Rochester, NY) filter apparatus and concentrated in Amicon (Danvers, MA) Centriprep 30 and Centricon 30 devices to a volume of less than 1.0 mL.

The concentrated periplasmic shockates from either the p49LHLH or p49LHHL clones were injected onto a Pharmacia (Piscataway, NJ) Superdex 75 HR 10/30 HPLC column that had been equilibrated with PBS. At a flow rate of 0.5 mL/minute, the product of interest, as determined by competition ELISA, had emerged between 21 through 24 minutes. The active fractions were pooled, concentrated as before and dialyzed overnight using a system 500 Microdialyzer Unit (Pierce Chemical) against 20 mM Tris-HCl pH 7.6 with 3-4 changes of buffer and using an 8,000 MW cut-off membrane. The sample was injected on a Pharmacia Mono Q HR 5/5 anion exchange HPLC column. A gradient program using 20 mM Tris-HCl pH 7.6 as buffer A and the same solution plus 0.5 M NaCl as buffer B was employed at a flow rate of 1.5 mL/min. The products of interest in each case, as determined by competition ELISA, emerged from the column between 3 and 4 minutes. Analysis of the fractions at this point on duplicate SDS-PAGE gels, one stained with Coomassie Brilliant Blue R-250 and the other transferred for Western analysis (using biotinylated FAID 14 as the probe antibody) revealed a single band at the calculated molecular weight for the scFv2 (LHLH or LHHL) species at 58,239 daltons. The active fractions were in each case concentrated, dialysed against 50 mM MES pH 5.8 overnight and injected on a Pharmacia Mono S HR 5/5 cation exchange column. The two fractions of interest from this purification step, as determined by SDS-PAGE and ELISA, fractions 5 and 6, eluted just before the start of the gradient, so they had not actually bound to the column. Fractions 5 and 6 were consequently pooled for future purification.

A Mono Q column was again run on the active Mono S fractions but the buffer used was 20 mM Tris-HCl, pH 8.0 and the flow rate was decreased to 0.8 mL/minute. The products emerged without binding, but the impurity left over from the Mono S was slightly more held up, so that separation did occur between 5 and 6 minutes. After this run, the products were homogeneous and were saved for further characterization.

Isoelectric Focusing

The isoelectric points (pI) of the constructs was predicted using the DNASTAR (Madison, WI) computer program Protein-titrate. Based on amino acid composition, a MW and pI value was calculated.

Experimentally, pIs were determined using FMC Bioproducts (Rockland, ME) Isogel IEF plates, pH range 3-10. A Biorad (Richmond, CA) electrophoresis unit was used to run the IEF, following the directions of both manufacturers. The electrophoresis conditions were as follows: 500 V (limiting) at 20 mA and at 10 W of constant power. Focusing was complete in 90 minutes. Biorad IEF standards included phycocyanin, beta lactoglobulin B, bovine carbonic anhydrase, human carbonic anhydrase, equine myoglobin, human hemoglobins A and C, 3 lentil lectin, and cytochrome C with pI value of 4.65, 5.10, 6.00, 6.50, 7.00, 7.50, 7.8, 8.00, 8.20

and 9.6, respectively. Gels were stained and destained according to directions provided by FMC. The DNASTAR program predicted values of 8.1 for the pI for both scFv2 species. A single, homogeneous band for the pure products was observed on the gel at pI values for both at 6.9.

Purified CC49 antibodies such as the IgG, scFv2 (LHLH and LHHL) were quantitated by measuring the absorbance spectrophotometrically at 280 nm. Molar absorptivity values, ϵ_M , were determined for each using the formula cited above by Wetlaufer.

Based on the amino acid composition, the $E^{0.1\%}$ (280 nanometers) values for CC49 IgG, CC49 scFv2 LHLH, CC49 scFv2 LHHL and CC49 scFv were 1.49, 1.65, 1.65 and 1.71, respectively.

10 Example 4

Relative activities of the CC49 scFv2 species LHLH and LHHL, were compared with the IgG and a monomer scFv form with a FLAG peptide at the COOH terminus.

Percent competition was determined from the ELISA data by the following equation:

$$15 \quad \frac{\text{Zero competition - sample reading (OD}_{405-450 \text{ nm}})}{\text{zero competition - 100 percent competition}} \times 100$$

The "zero competition" value was determined by mixing (1:1) one percent BSA with the biotinylated CC49 (3×10^{-14} moles) while the 100 percent competition value was based on a 5 $\mu\text{g/mL}$ sample of CC49 IgG mixed with the biotinylated CC49 IgG. The data are presented in Figure 11. Absorbance values for the samples were measured at 405 nm - 450 nm.

20 The average of triplicate readings was used. Initially samples (25 μL) were applied to the TAG-72 coated microliter plates at 1.0×10^{-10} moles of binding sites/mL. Biotinylated CC49 (4 $\mu\text{g}/\mu\text{L}$ diluted 1:20,000 - used 25 μL) diluted the samples by a factor of 2. Serial dilutions (1:2) were performed. Both forms of the scFv2 are approximately equivalent to the IgG (see Figure 11). In a separate experiment, a CC49 scFv monomer was compared to a Fab fragment,

25 both of which are monovalent and these were also shown to be equivalent in their binding affinity for TAG-72. These results indicate that both forms of the covalently linked dimers have 2 fully functional antigen binding sites. This is the same increase in avidity as observed with the whole IgG, relative to a monomeric species.

These data also indicate that the scFv2 molecules, like their CC49 IgG parent are

30 candidates for immunotherapeutic applications, but with the benefit of increased capillary permeability and more rapid biodistribution pharmacokinetics. The advantage should allow multiple injections of compounds of the present invention and give higher tumor:tissue ratios in immunotherapeutic treatment regimens for cancer treatment, relative to the existing IgG molecules.

35 Other embodiments of the invention will be apparent to those skilled in the art from a consideration of this specification or practice of the invention disclosed herein. It is

intended that the specification and examples be considered as exemplary only, with the true scope and spirit of the invention being indicated by the following claims.

5

10

15

20

25

30

35

1. A multivalent single chain antibody which comprises two or more single chain antibody fragments each fragment having affinity for an antigen wherein the fragments are covalently linked by a first peptide linker and each fragment comprising:
- 5 (a) a first polypeptide comprising a light chain variable domain;
(b) a second polypeptide comprising a heavy chain variable domain; and
(c) a second peptide linker linking the first and second polypeptides into a functional binding moiety.
2. The multivalent single chain antibody of Claim 1 wherein the first peptide linker has the amino acid sequence
- 10 Leu Ser Ala Asp Asp Ala Lys Lys Asp Ala Ala Lys Lys Asp Asp Ala Lys Lys Asp Asp Ala Lys Lys Asp Leu.
3. The multivalent single chain antibody of Claim 1 wherein the light chain variable region has an amino acid sequence substantially the same as that of Figure 3 and the heavy chain variable region has an amino acid sequence substantially the same as that of
- 15 Figure 5.
4. The multivalent single chain antibody of Claim 1 wherein the first and second peptide linkers have an amino acid sequence which is substantially the same.
5. A DNA sequence which codes for a multivalent single chain antibody, the multivalent single antibody comprising two or more single chain antibody fragments, each
- 20 fragment having affinity for an antigen wherein the fragments are covalently linked by a first peptide linker and each fragment comprising:
- (a) a first polypeptide comprising a light chain variable domain;
(b) a second polypeptide comprising a heavy chain variable domain; and
(c) a second peptide linker linking the first and second polypeptides into a functional binding moiety.
- 25
6. The DNA sequence of Claim 5 wherein the sequence coding for the first polypeptide is substantially the same as that of Figure 2 and the sequence coding for the second polypeptide is substantially the same as that of Figure 3.

30

35

FIG. 1

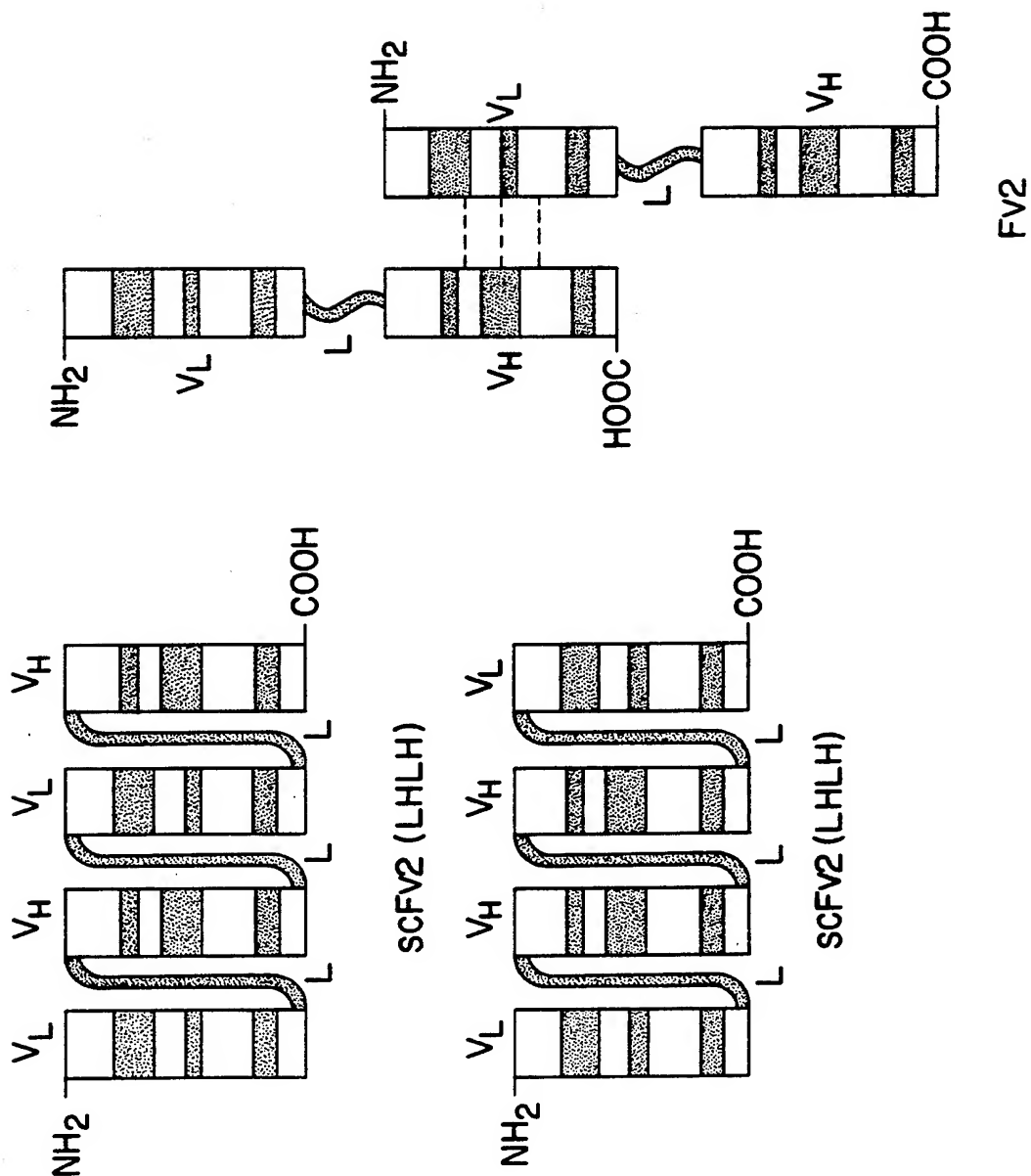


FIG. 2

GAC ATT GTG ATG TCA CAG TCT CCA TCC TCC CTA CCT GTG TCA
 GTT GGC GAG AAG GTT ACT TTG AGC TGC AAG TCC AGT CAG AGC
 CTT TTA TAT AGT GGT AAT CAA AAG AAC TAC TTG GCC TGG TAC
 CAG CAG AAA CCA GGG CAG TCT CCT AAA CTG CTG ATT TAC TGG
 GCA TCC GCT AGG GAA TCT GGG GTC CCT GAT CGC TTC ACA GGC
 AGT GGA TCT GGG ACA GAT TTC ACT CTC TCC ATC AGC AGT GTG
 AAG ACT GAA GAC CTG GCA GTT TAT TAC TGT CAG CAG TAT TAT
 AGC TAT CCC CTC ACG TTC GGT GCT GGG ACC AAG CTG GTG CTG
 AAG

FIG. 3

Asp Ile Val Met Ser Gln Ser Pro Ser Ser Leu Pro Val Ser Val
 Gly Glu Lys Val Thr Leu Ser Cys Lys Ser Ser Gln Ser Leu Leu
 Tyr Ser Gly Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys
 Pro Gly Gln Ser Pro Lys Leu Leu Ile Tyr Trp Ala Ser Ala Arg
 Glu Ser Gly Val Pro Asp Arg Phe Thr Gly Ser Gly Ser Gly Thr
 Asp Phe Thr Leu Ser Ile Ser Ser Val Lys Thr Glu Asp Leu Ala
 Val Tyr Tyr Cys Gln Gln Tyr Tyr Ser Tyr Pro Leu Thr Phe Gly
 Ala Gly Thr Lys Leu Val Leu Lys

FIG. 4

GAG GTT CAG TTG CAG CAG TCT GAC GCT GAG TTG GTG AAA CCT
GGG GCT TCA GTG AAG ATT TCC TGC AAG GCT TCT GGC TAC ACC
TTC ACT GAC CAT GCA ATT CAC TGG GTG AAA CAG AAC CCT GAA
CAG GGC CTG GAA TGG ATT GGA TAT TTT TCT CCC GGA AAT GAT
GAT TTT AAA TAC AAT GAG AGG TTC AAG GGC AAG GCC ACA CTG
ACT GCA GAC AAA TCC TCC AGC ACT GCC TAC GTG CAG CTC AAC
AGC CTG ACA TCT GAG GAT TCT GCA GTG TAT TTC TGT ACA AGA
TCC CTG AAT ATG GCC TAC TGG GGT CAA GGA ACC TCA GTC ACC
GTC TCC TCA

FIG. 5

Glu Val Gln Leu Gln Gln Ser Asp Ala Glu Leu Val Lys Pro Gly
Ala Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr
Asp His Ala Ile His Trp Val Lys Gln Asn Pro Glu Gln Gly Leu
Glu Trp Ile Gly Tyr Phe Ser Pro Gly Asn Asp Asp Phe Lys Tyr
Asn Glu Arg Phe Lys Gly Lys Ala Thr Leu Thr Ala Asp Lys Ser
Ser Ser Thr Ala Tyr Val Gln Leu Asn Ser Leu Thr Ser Glu Asp
Ser Ala Val Tyr Phe Cys Thr Arg Ser Leu Asn Met Ala Tyr Trp
Gly Gln Gly Thr Ser Val Thr Val Ser Ser

FIG. 6A

DNA AND AMINO ACID SEQUENCE OF CC49 VL-L-VH-L-VL-L-VH

| | | | | | | | | | | | | | | | | |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 5'-C | TCA | TGT | TTG | ACA | GCT | TAT | CAT | CGA | TGA | ATT | CCA | TCA | CTT | CCC | TCC | 46 |
| | | | | | | | | | | | | | | | | |
| GTT | CAT | TTG | TCC | CCG | GTG | GAA | ACG | AGG | TCA | TCA | TTT | CCT | TCC | GAA | AAA | 94 |
| ACG | GTT | GCA | TTT | AAA | TCT | TAC | ATA | TAT | AAT | ACT | TTC | AAA | GAC | TAC | ATT | 142 |
| TGT | AAG | ATT | TGA | TGT | TTG | AGT | CGG | CTG | AAA | GAT | CGT | ACG | TAC | CAA | TTA | 190 |
| TTG | TTT | CGT | GAT | TGT | TCA | AGC | CAT | AAC | ACT | GTA | GGG | ATA | GTG | GAA | AGA | 238 |
| GTG | CTT | CAT | CTG | GTT | ACG | ATC | AAT | CAA | ATA | TTC | AAA | CGG | AGG | GAG | ACG | 286 |
| | | | | | | | | | | | | | | | | |
| ATT | TTG | ATG | AAA | TAC | CTA | TTG | CCT | ACG | GCA | GCC | GCT | GGA | TTG | TTA | TTA | 334 |
| | | | | | | | | | | | | | | | | |
| Leu | Ala | Ala | Gln | Pro | Ala | Met | Ala | Asp | Ile | Val | Met | Ser | Gln | Ser | Pro | |
| CTC | GCT | GCC | CAA | CCA | GCC | ATG | GCC | GAC | ATT | GTG | ATG | TCA | CAG | TCT | CCA | 382 |
| | | | | | | | | | | | | | | | | |
| Ser | Ser | Leu | Pro | Val | Ser | Val | Gly | Glu | Lys | Val | Thr | Leu | Ser | Cys | Lys | |
| TCC | TCC | CTA | CCT | GTG | TCA | GTT | GGC | GAG | AAG | GTT | ACT | TTG | AGC | TGC | AAG | 430 |
| | | | | | | | | | | | | | | | | |
| Ser | Ser | Gln | Ser | Leu | Leu | Tyr | Ser | Gly | Asn | Gln | Lys | Asn | Tyr | Leu | Ala | |
| TCC | AGT | CAG | AGC | CTT | TTA | TAT | AGT | GGT | AAT | CAA | AAG | AAC | TAC | TTG | GCC | 478 |

FIG. 6B

| | |
|---|-----|
| Trp Tyr Gln Gln Lys Pro Gly Gln Ser Pro Lys Leu Leu Ile Tyr Trp | 526 |
| TGG TAC CAG CAG AAA CCA CCA GGG CAG TCT CCT AAA CTG CTG ATT TAC TGG | |
| Ala Ser Ala Arg Gln Ser Gly Val Pro Asp Arg Phe Thr Gly Ser Gly | 574 |
| GCA TCC GCT AGG GAA TCT GGG GTC CCT GAT CGC TTC ACA GGC AGT GGA | |
| Ser Gly Thr Asp Phe Thr Leu Ser Ile Ser Ser Val Lys Thr Glu Asp | 622 |
| TCT GGG ACA GAT TTC ACT CTC TCC ATC AGC AGT GTG AAG ACT GAA GAC | |
| Leu Ala Val Tyr Tyr Cys Gln Gln Tyr Tyr Ser Tyr Pro Leu Thr Phe | 670 |
| CTG GCA GTT TAT TAC TGT CAG CAG TAT TAT AGC TAT CCC CTC ACG TTC | |
| Gly Ala Gly Thr Lys Lys Leu Val Leu Lys Leu Ser Ala Asp Asp Ala Lys | 718 |
| GGT GCT GGG ACC AAG CTG GTG CTG AAG CTT AGT GCG GAC GAT GCG AAA | |
| Lys Asp Ala Ala Lys Lys Lys Asp Asp Ala Lys Lys Asp Asp Ala Lys Lys | 766 |
| AAG GAT GCT GCG AAG AAG AAG GAT GAC GCT AAG GAT GAC GAT GCT AAA AAG | |

FIG. 6C

[illegible]

FIG. 6D

| | | | | |
|---|---|--|--|------|
| 240 | | VH49J- G AAT ATG GCC TAC TGG GGT CAA G | | |
| Phe | Cys Thr Arg Ser Leu Asn Met Ala Tyr Trp Gly Gln Gly Thr Ser | | | |
| TTC TGT ACA AGA TCC CTG AAT ATG GCC TAC TGG GGT CAA GGA ACC TCA | | | | 1102 |
| 250 | | 260 | | |
| Val Thr Val Ser Ser Leu Ser Ala Asp Asp Ala Lys Lys Asp Ala Ala | | | | |
| GTC ACC GTC TCC TCA CTA AGC GCA GAT GAC GCA AAG AAA GAC GCA GCT | | | | 1150 |
| 270 | | VL 280 | | |
| Lys Lys Asp Asp Ala Lys Lys Asp Ala Lys Lys Asp Leu Asp Ile | | | | |
| AAA AAA GAC GAT GCC AAA AAG GAT GAC GCC AAG AAA GAT CTT GAC ATT | | | | 1198 |
| 290 | | 1246 | | |
| Val Met Ser Gln Ser Pro Ser Ser Leu Pro Val Ser Val Gly Glu Lys | | | | |
| GTG ATG TCA CAG TCT CCA TCC TCC CTA CCT GTG TCA GTT GGC GAG AAG | | | | |
| 300 | | 310 | | |
| Val Thr Leu Ser Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser Gly Asn | | | | |
| GTT ACT TTG AGC TGC AAG TCC AGT CAG AGC CTT TTA TAT AGT GGT AAT | | | | 1294 |
| 320 | | 1342 | | |
| Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ser Pro | | | | |
| CAA AAG AAC TAC TTG GCC TGG TAC CAG CAG AAA CCA GGC CAG TCT CCT | | | | |

FIG. 6E

| | | | |
|---|---|------------|------|
| 330 | Lys Leu Leu Ile Tyr Trp Ala Ser Ala Arg Glu Ser Gly Val Pro Asp | 340 | 1390 |
| AAA CTG CTG ATT TAC TGG GCA TCC GCT AGG GAA TCT GGG GTC CCT GAT | | | |
| Arg Phe Thr Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Ser Ile Ser | 350 | 360 | |
| CGC TTC ACA GGC AGT GGA TCT GGG ACA GAT TTC ACT CTC TCC ATC AGC | | | 1438 |
| Ser Val Lys Thr Glu Asp Leu Ala Val Tyr Tyr Cys Gln Gln Tyr Tyr | 370 | 49VLCDR3+- | 1486 |
| AGT GTG AAG ACT GAA GAC CTG GCA GTT TAT TAC TGT CAG CAG TAT TAT | | | |
| AGC TAT | 380 | 390 | 1534 |
| Ser Tyr Pro Leu Thr Phe Gly Ala Gly Thr Lys Leu Val Leu Lys Leu | | | |
| AGC TAT CCC CTC ACG TTC GGT GGT GCT GGG ACC AAG CTG GTG CTG AAG CTA | | | |
| Eco47 III | 400 | | 1582 |
| Ser Ala Asp Asp Ala Lys Lys Lys Lys Lys Lys Asp Asp Ala Lys | | | |
| AGC GCT GAT GAT GAT GCT AAG AAG AAG AAG AAG AAG AAG GAC GCA AAA | | | |
| Lys Asp Asp Ala Lys Lys Asp Leu Glu Val Gln Leu Gln Gln Ser Asp | 410 | 420 | 1630 |
| AAG GAT GAT GAT GCA AAA AAG AAG AAG GAT CTG GAG GTT CAG TTG CAG CAG TCT GAC | | | |
| Ala Glu Leu Val Lys Pro Gly Ala Ser Val Lys Ile Ser Cys Lys Ala | 430 | 440 | 1678 |
| GCT GAG TTG GTG AAA CCT GGG GCT TCA GTG AAG ATT TCC TGC AAG GCT | | | |

FIG. 6F

| | | |
|---|-------------------------|------|
| Ser Gly Tyr Thr Phe Thr Asp His Ala Ile | His Trp Val Lys Gln Asn | 1726 |
| TCT GGC TAC ACC TTC ACT GAC CAT GCA ATT | CAC TGG GTG AAA CAG AAC | |
| Pro Glu Gln Gly Leu Glu Trp Ile Gly Tyr Phe Ser Pro Gly Asn Asp | | 1774 |
| CCT GAA CAG GGC CTG GAA TGG ATT GGA TAT TTT TCT CCC GGA AAT GAT | | |
| Asp Phe Lys Tyr Asn Glu Glu Arg Phe Lys Gly Lys Ala Thr Leu Thr Ala | | 1822 |
| GAT TTT AAA TAC AAT GAG AGG TTC AAG GGC AAG GCC ACA CTG ACT GCA | | |
| Asp Lys Ser Ser Thr Ala Tyr Val Gln Leu Asn Ser Leu Thr Ser | | 1870 |
| GAC AAA TCC TCC AGC ACT GCC TAC TAC GTG CAG CTC AAC AGC CTG ACA TCT | | |
| Glu Asp Ser Ala Val Tyr Phe Cys Thr Arg Ser Leu Asn Met Ala Tyr | | 1918 |
| GAG GAT TCT GCA GTG TAT TTC TGT ACA AGA TCC CTG AAT ATG GCC TAC | | |
| Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser | Nhe I | 1966 |
| TGG GGT CAA GGA ACC ACC GTC ACC GTC TCC TCA TAA AAA GCT AGC GAT | | |

FIG. 6G

| | |
|---|------|
| GAA TCC GTC AAA ACA TCA TCT TAC ATA AAG TCA CTT GGT GAT CAA GCT | 2014 |
| SQP1- TGT AGT AGA ATG TAT TTC AGT | |
| PENPTSEQ2- G TAT TTC AGT GAA CCA CTA GTT | |
| CAT ATC ATT GTC CGG CAA TGG TGT GGG CTT TTT TTG TTT TCT ATC TTT | 2062 |
| AAA GAT CAT GTG AAG AAA AAC GGG AAA ATC GGT CTG CGG GAA AGG ACC | 2110 |
| GGG TTT TTG TCG AAA TCA TAG GCG AAT GGG TTG GAT TGT GAC AAA ATT | 2158 |
| BamH I | |
| CGG ATC C-3' | 2165 |

FIG. 7A

DNA AND AMINO ACID SEQUENCE OF CC49 VL-L-VH-L-VH-L-VL

| | |
|--|-----|
| 5'-C TCA TGT TTG ACA GCT TAT CAT CGA TGA ATT CCA TCA CTT CCC TCC | 46 |
| GTT CAT TTG TCC CCG GTG GAA ACG AGG TCA TCA TTT CCT TCC GAA AAA | 94 |
| ACG GTT GCA TTT AAA TCT TAC ATA TAT AAT ACT TTC AAA GAC TAC ATT | 142 |
| TGT AAG ATT TGA TGT TTG AGT CGG CTG AAA GAT CGT ACG TAC CAA TTA | 190 |
| TTG TTT CGT GAT TGT TCA AGC CAT AAC ACT GTA GGG ATA GTG GAA AGA | 238 |
| GTG CTT CAT CTG GTT ACG ATC AAT CAA ATA TTC AAA CGG AGG GAG ACG | 286 |
| PENPR2- TAT AAG TTT GCC TCC CTC TG | |
| -22 Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu | 334 |
| ATT TTG ATG AAA TAC CTA TTG CCT ACG GCA GCC GCT GGA TTG TTA TTA | |
| Nco I VL | |
| Leu Ala Ala Gln Pro Ala Met Ala Asp Ile Val Met Ser Gln Ser Pro | 382 |
| CTC GCT GCC CAA CCA GCC ATG GCC GAC ATT GTG ATG TCA CAG TCT CCA | |
| 10 Ser Ser Leu Pro Val Ser Val Gly Glu Lys Val Thr Leu Ser Cys Lys | 430 |
| TCC TCC CTA CCT GTG TCA GTT GGC GAG AAG GTT ACT TTG AGC TGC AAG | |

FIG. 7B

| | |
|---|-----|
| Ser Ser Gln Ser Leu Tyr Ser Gly Asn Gln Lys Asn Tyr Leu Ala | 40 |
| TCC AGT CAG AGC CTT TTA TAT AGT GGT AAT CAA AAG AAC TAC TTG GCC | 478 |
| Trp Tyr Gln Gln Lys Pro Gly Gln Ser Pro Lys Leu Ile Tyr Trp | |
| TGG TAC CAG CAG AAA CCA CCA GGG CAG TCT GCT AAA CTG CTG ATT TAC TGG | 526 |
| Ala Ser Ala Arg Glu Ser Gly Val Pro Asp Arg Phe Thr Gly Ser Gly | 70 |
| GCA TCC GCT AGG GAA TCT GGG GTC CCT GAT CGC TTC ACA GGC AGT GGA | 574 |
| Ser Gly Thr Asp Phe Thr Leu Ser Ile Ser Ser Val Lys Thr Glu Asp | |
| TCT GGG ACA GAT TTC ACT CTC TCC TCC ATC AGC AGT GTG AAG ACT GAA GAC | 622 |
| Leu Ala Val Tyr Tyr Cys Gln Gln Tyr Tyr Ser Tyr Pro Leu Thr Phe | 100 |
| CTG GCA GTT TAT TAC TGT CAG CAG TAT TAT AGC TAT CCC CTC ACG TTC | 670 |
| Gly Ala Gly Thr Lys Leu Val Leu Lys Leu Ser Ala Asp Asp Ala Lys | 120 |
| GGT GCT GGG ACC AAG CTG GTG CTG AAG CTT AGT GCG GAC GAT GCG AAA | 718 |
| Lys Asp Ala Ala Lys Lys Lys Asp Asp Ala Lys Lys Lys Lys Lys | 130 |
| AAG GAT GCT GCG AAG AAG GAT GAC GCT AAG AAA GAC GAT GCT AAA AAG | 766 |
| TTC CTA CGA CGC TTC TTC CTA TMNVL(-)SEQ | |

FIG. 7C

[illegible]

FIG. 7D

| | | | | |
|-----|---|-----|---|------|
| 250 | Val Thr Val Ser Ser Leu Ser Ala Asp Asp Ala Lys Lys Asp Ala Ala | 260 | Lys Lys Asp Ala Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys | 1150 |
| | GTC ACC GTC TCC TCA CTA AGC GCA GAT GAC GCA AAG AAA GAC GCA GCT | | VH 280 | |
| | | 270 | Lys Lys Asp Asp Ala Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys | 1198 |
| | | | AAA AAA GAC GAT GCC AAA AAG AAG GAT GAC GCC AAG AAA GAT CTT GAG GTT | |
| | | 290 | Gln Leu Gln Gln Ser Asp Ala Gln Leu Val Lys Pro Gly Ala Ser Val | 1246 |
| | | | CAG TTG CAG CAG CAG TCT GAC GCT GAG TTG GTG AAA CCT GGG GCT TCA GTG | |
| | | 300 | Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asp His Ala Ile | 1294 |
| | | | AAG ATT TCC TGC AAG GCT TCT TCT GGC TAC ACC TTC ACT GAC CAT GCA ATT | |
| | | 320 | His Trp Val Lys Gln Asn Pro Gln Gly Leu Glu Trp Ile Gly Tyr | 1342 |
| | | | CAC TGG GTG AAA CAG AAC CCT GAA CAG GGC CTG GAA TGG ATT GGA TAT | |
| | | 330 | Phe Ser Pro Gly Asn Asp Phe Lys Tyr Asn Glu Arg Phe Lys Gly | 1390 |
| | | | TTT TCT CCC GGA AAT GAT GAT TTT AAA TAC AAT GAG AGG TTC AAG GGC | |
| | | 350 | Lys Ala Thr Leu Thr Ala Asp Lys Ser Ser Ser Thr Ala Tyr Val Gln | 1438 |
| | | | AAG GCC ACA CTG ACT GCA GAC AAA TCC TCC AGC ACT GCC TAC GTG CAG | |
| | | 370 | Leu Asn Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys Thr Arg | 1486 |
| | | | CTC AAC AGC CTG ACA TCT GAG GAT TCT GCA GTG TAT TTC TGT ACA AGA | |

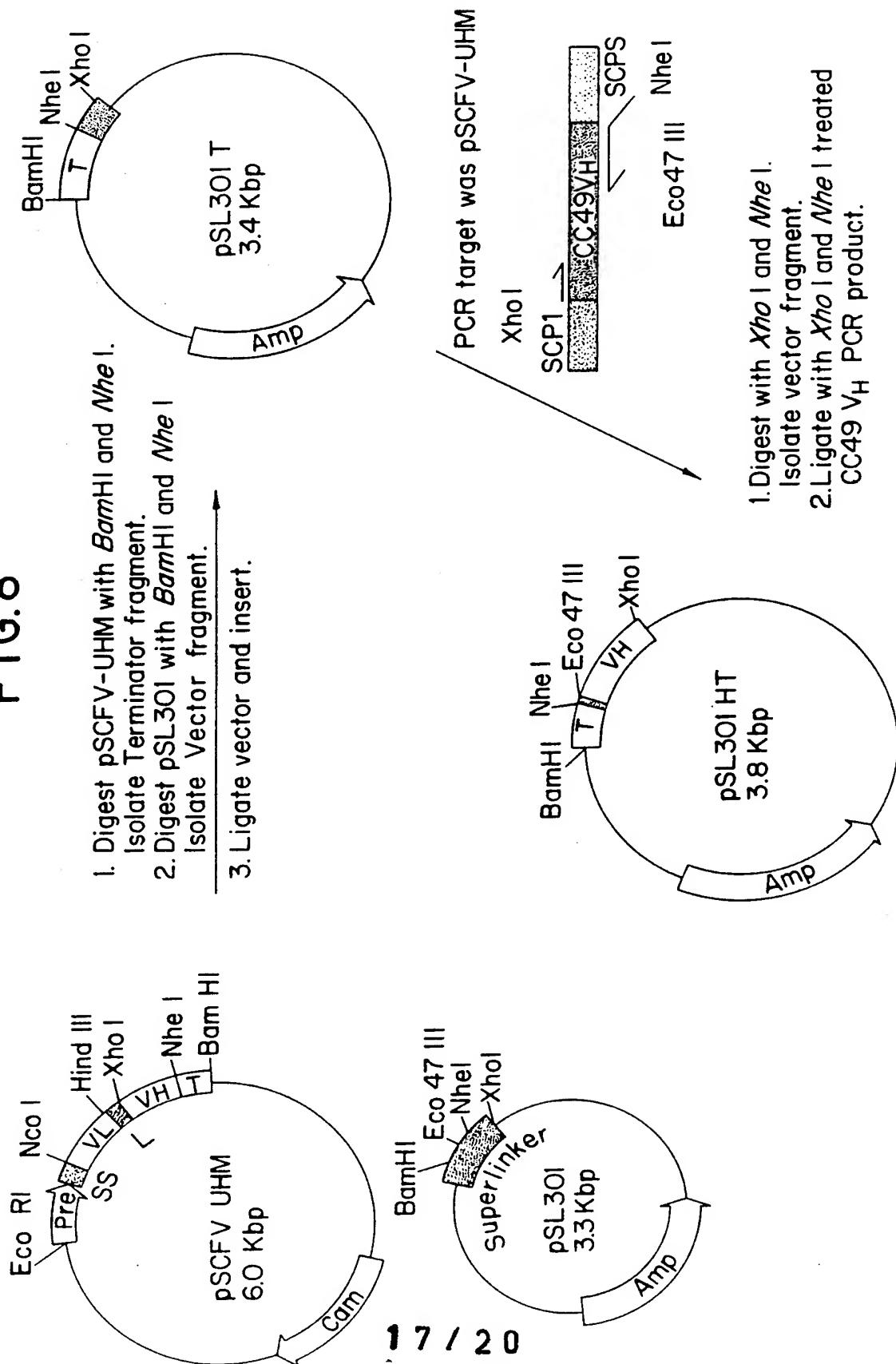
FIG. 7F

| | | | |
|-----|---|-----|------|
| 490 | Ser Gly Ser Gly Thr Asp Phe Thr Leu Ser Ile Ser Ser Val Lys Thr | 500 | 1870 |
| | AGT GGA TCT GGG ACA GAT TTC ACT CTC TCC ATC AGC AGT AGT GTG AAG ACT | | |
| | | | |
| | 510 | | |
| | Glu Asp Leu Ala Val Tyr Cys Gln Gln Tyr Tyr Ser Tyr Pro Leu | 520 | |
| | GAA GAC CTG GCA GTT TAT TAC TGT CAG CAG TAT TAT AGC TAT CCC CTC | | 1918 |
| | | | |
| | 530 | | |
| | Thr Phe Gly Ala Gly Thr Lys Leu Val Leu Lys *** Nhe I | | 1966 |
| | ACG TTC GGT GCT GGG ACC AAG CTG CTG GTG CTT AAG TAA AAA GCT AGC GAT | | |
| | | | |
| | GAA TCC GTC AAA ACA TCA TCT TAC ATA AAG TCA CTT GGT GAT CAA GCT | | 2014 |
| | SQP1- TGT AGT AGA ATG TAT TTC AGT | | |
| | PENPTSEQ2- G TAT TTC AGT GAA CCA CTA GTT | | |
| | | | |
| | CAT ATC ATT GTC CGG CAA TGG TGT GGT CTT TTT TTG TTT TCT ATC TTT | | 2062 |
| | | | |
| | AAA GAT CAT GTG AAG AAA AAC GGG AAA ATC GGT CTG CGG GAA AGG ACC | | 2110 |
| | | | |
| | GGG TTT TTG TCG AAA TCA TAG GCG AAT GGG TTG GAT TGT GAC AAA ATT | | 2158 |
| | | | |
| | BamH I | | |
| | CGG ATC C-3' | | 2165 |

FIG. 7E

| | | | |
|--|-----|---------------|------|
| Ser Leu Asn Met Ala Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser TCC CTG AAT ATG GCC TAC TGG GGT CAA GGA ACC TCA GTC ACC GTC TCC | 380 | 390 | 1534 |
| Ser Leu Ser Ala Asp Asp Ala Lys Lys Asp Ala Ala Lys Lys Asp Asp TCA CTA AGC GCA GAT GAC GAC AAG AAA GAC GCA GCT AAA AAA GAC GAT | 400 | | 1582 |
| Ala Lys Lys Asp Asp Ala Lys Lys Asp Leu Asp Ile Val Met Ser Gln GCC AAA AAG GAT GAC GCC AAG AAA GAT CTT GAC ATT GTG ATG TCA CAG | 410 | VL 420 | 1630 |
| Ser Pro Ser Ser Leu Pro Val Ser Val Gly Glu Lys Val Thr Leu Ser TCT CCA TCC TCC CTA CCT GTG TCA GTT GGC GAG AAG GTT ACT TTG AGC | 430 | 440 | 1678 |
| Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser Gly Asn Gln Lys Asn Tyr TGC AAG TCC AGT CAG AGC CTT TTA TAT AGT GGT AAT CAA AAG AAC TAC | 450 | 49LFR2(-) - G | 1726 |
| Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ser Pro Lys Leu Leu Ile TTG GCC TGG TAC CAG CAG CCA GGG CAG TCT CCT CCT AAA CTG CTG ATT | 460 | 470 | 1774 |
| Tyr Trp Ala Ser Ala Arg Glu Ser Gly Val Pro Asp Arg Phe Thr Gly TAC TGG GCA TCC GCT AGG GAA TCT TCT GGG GTC CCT GAT CGC TTC ACA GGC | 480 | | 1822 |

FIG. 8



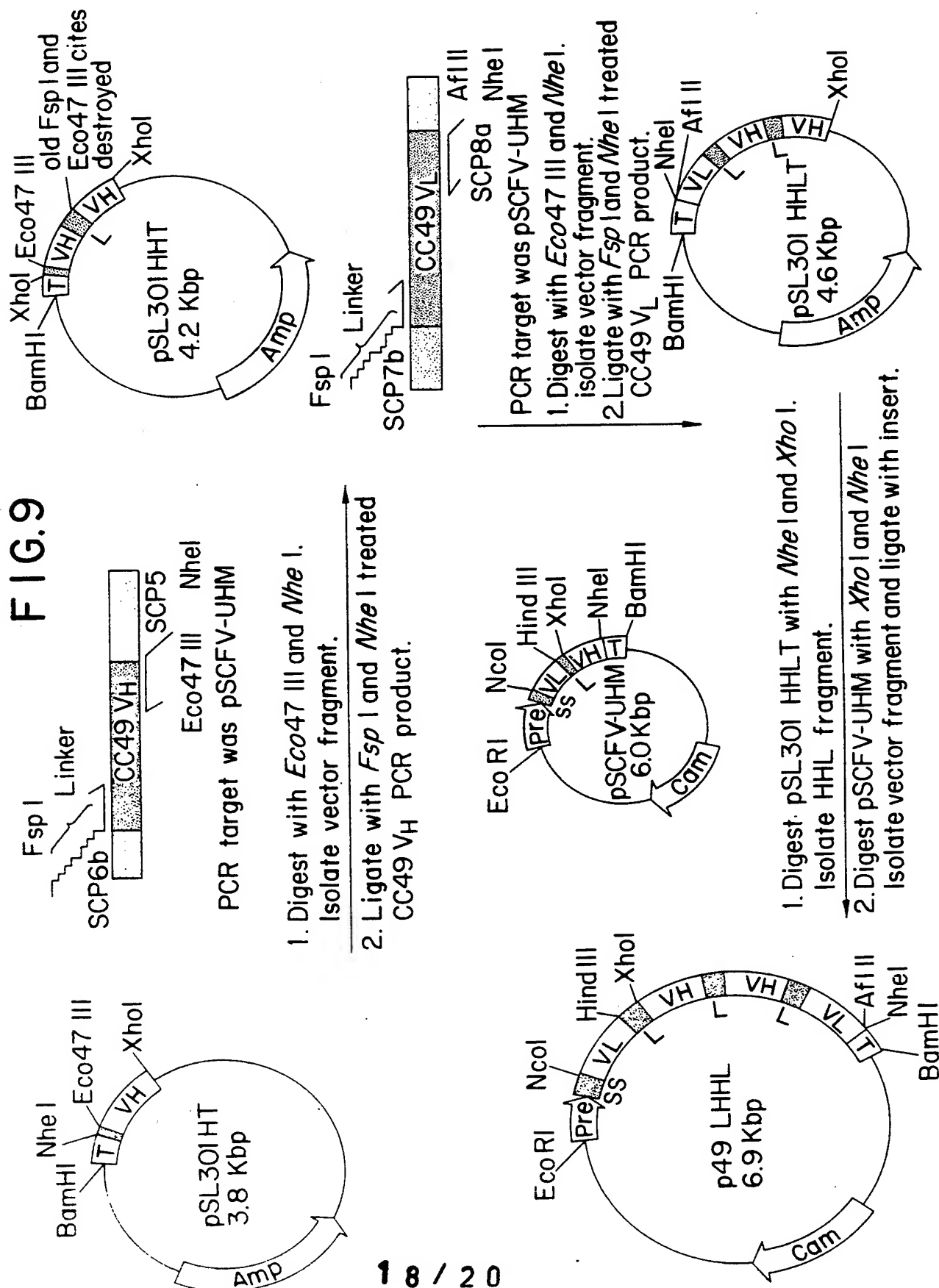


FIG.10

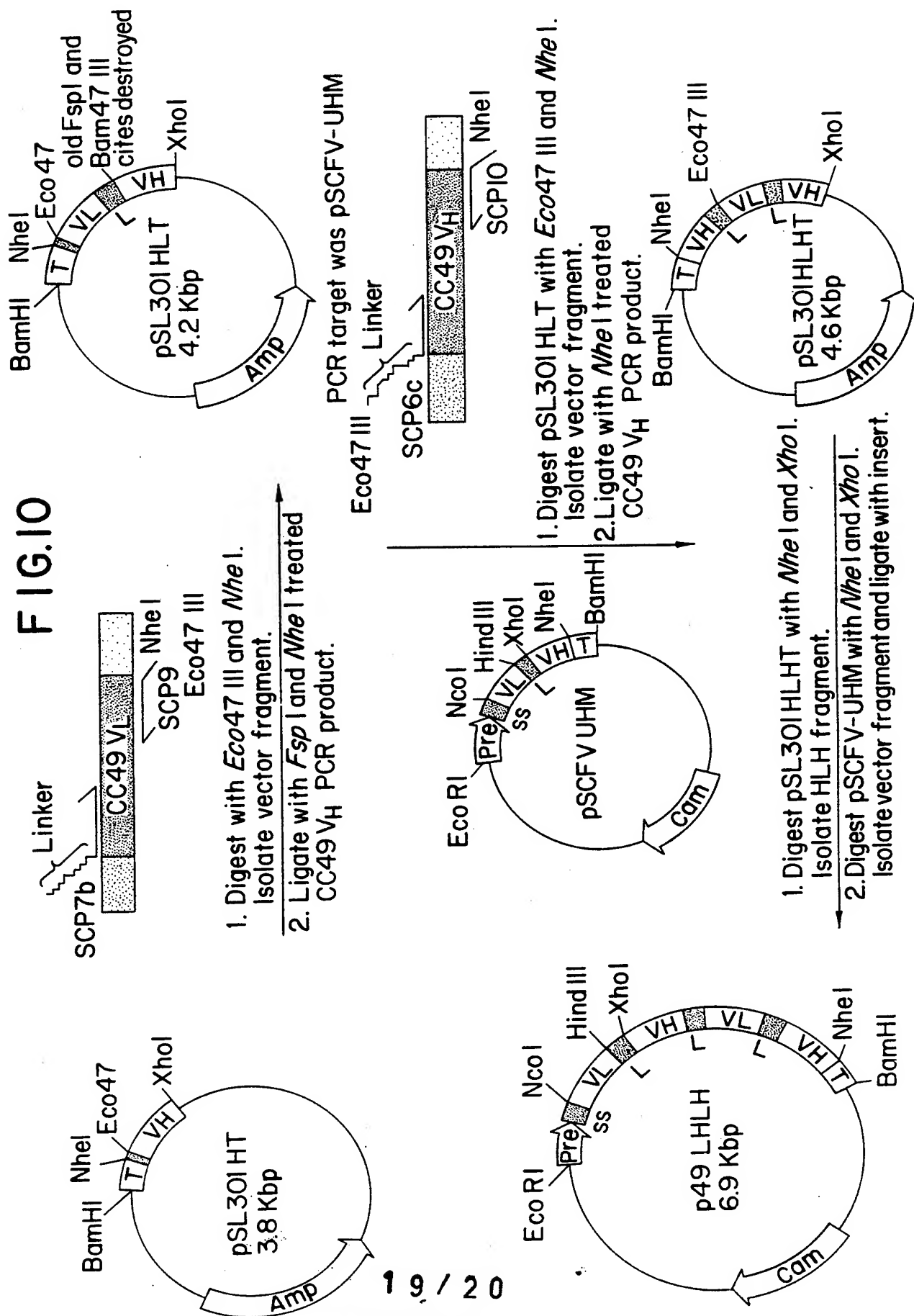
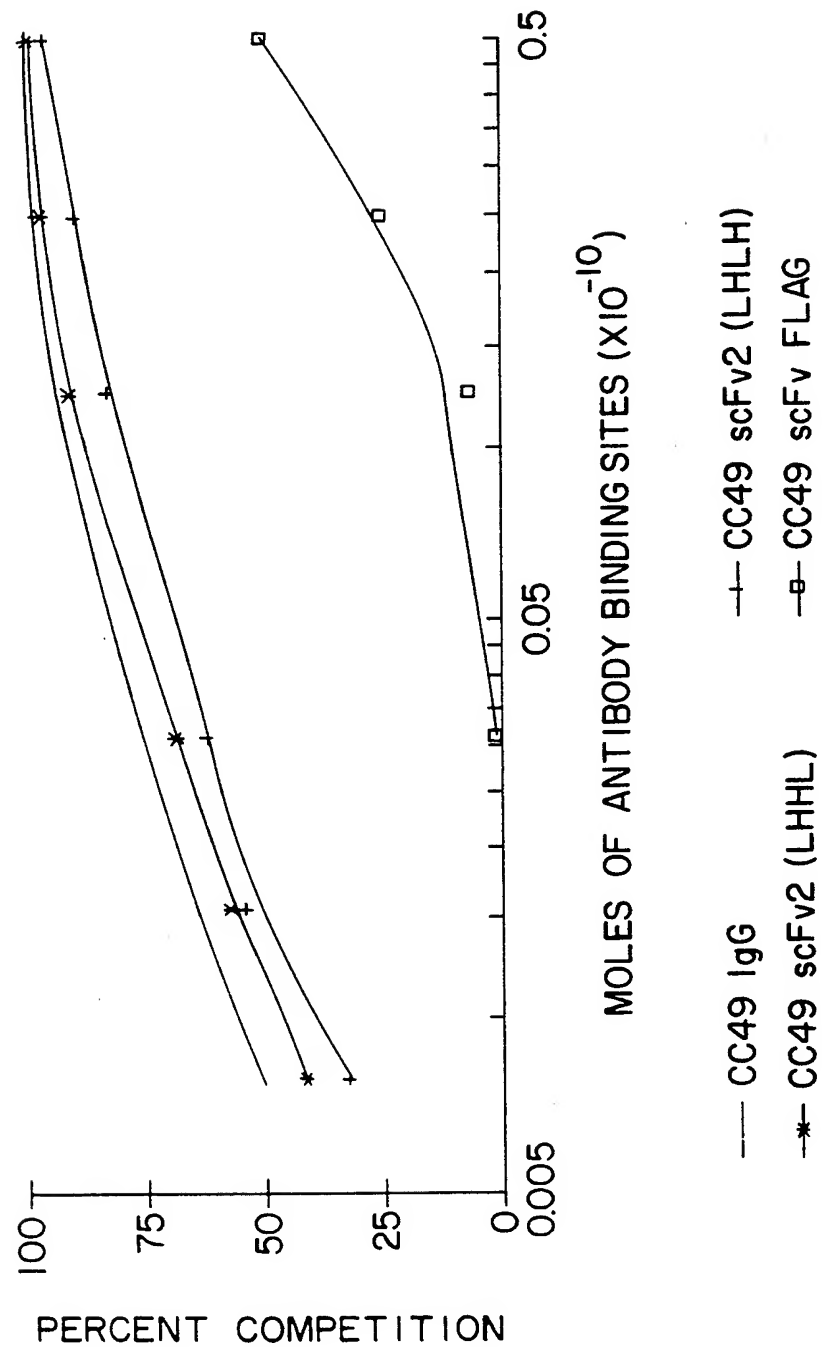


FIG. II



INTERNATIONAL SEARCH REPORT

Intern al Application No

PCT/US 93/12039

A. CLASSIFICATION OF SUBJECT MATTER

IPC 5 C12N15/13 C07K15/28 C12N15/62 A61K39/395

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 5 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|------------|--|-----------------------|
| X | WO,A,91 19739 (CELLTECH LIMITED) 26 December 1991 | 1,5 |
| Y | see example 1 --- | 2-4,6 |
| Y | CANCER RESEARCH vol. 52, no. 12 , 15 June 1992 , PHILADELPHIA, PA, USA pages 3402 - 3408 T.YOKATA ET AL. 'Rapid tumour penetration of a single-chain Fv and comparison with other immunoglobulin forms' see page 3403, column 1, paragraph 4 --- -/-- | 3,6 |

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

25 March 1994

Date of mailing of the international search report

27 -04- 1994

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Authorized officer

Cupido, M

INTERNATIONAL SEARCH REPORT

Intern: 11 Application No

PCT/US 93/12039

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|------------|---|-----------------------|
| Y | BIOCHEMISTRY vol. 30, no. 42 , 22 October 1991 , EASTON, PA US pages 10117 - 10125 M.W.PANTOLIANO ET AL. 'Conformational stability, folding and ligand-binding affinity of single-chain Fv immunoglobulin fragments expressed in Escherichia coli' cited in the application see page 10120, column 1, paragraph 2 --- | 2,4 |
| X | EP,A,0 506 124 (TANOX BIOSYSTEMS, INC.) 30 September 1992 see example 4 --- | 1,5 |
| P,X | WO,A,93 11161 (ENZON, INC.) 10 June 1993 see figure 19A ----- | 1,3-6 |

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 93/12039

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---------------------|--|--|
| WO-A-9119739 | 26-12-91 | AU-A- 7983191 EP-A- 0486652 GB-A- 2250995 JP-T- 5502039 | 07-01-92 27-05-92 24-06-92 15-04-93 |
| EP-A-0506124 | 30-09-92 | AU-B- 640863 AU-A- 1299292 JP-A- 5117164 | 02-09-93 15-10-92 14-05-93 |
| WO-A-9311161 | 10-06-93 | AU-A- 3178993 | 28-06-93 |